hexagon:2611.00368v2

Supersymmetric Analysis of a Random Band Matrix at Criticality

Volume I

Patrick Lopatto

AUTHOR'S NOTE

With the exception of this note, the entire document is machine-written. I have carefully supervised its preparation to ensure that the proofs are readable and that previous literature is appropriately cited. I welcome suggestions concerning mathematical corrections, omitted citations, and attribution of ideas.

The author was partially supported by NSF grant DMS-2450004.

ABSTRACT

We study the $N \times N$ complex Hermitian Gaussian band matrix whose variance profile is the inverse of $\mathrm{Id} - W^2\Delta$, where Δ is the Neumann Laplacian on $\{1, \dots, N\}$, in the critical regime $W^2 \asymp N$. We prove that, for energies E in a neighborhood of zero, the unfolded eigenvalue process converges to a stationary point process, whose law depends only on the limit X of $N/((4 - E^2)W^2)$ and is given by the partition functions of a supersymmetric sigma model on an interval. We also determine the macroscopic law of the mass of an eigenvector selected near a given energy. Its density is the normalized exponential of a Brownian motion with drift, under the law tilted by its total mass; the same law describes the eigenvectors of the critical one-dimensional random Schrödinger operator. Further, we compute the limits of the two-energy correlations of the local densities of states at all pairs of interior sites, of the mean eigenvector profile at the scale of single sites, and of N times the mean inverse participation ratio, which is $2 + 2X/3$. As $X \to 0$ and $X \to \infty$, we recover the pair statistics of the GUE and of a Poisson process, and in the localized limit we obtain an explicit random shape for the eigenvectors. The number variance in a window of length L grows like $\sqrt{XL}$ as $L \to \infty$ at every fixed X, and we describe the crossover to Poisson statistics in windows of length of order X. Several of these results confirm predictions that physicists obtained from the sigma model. The proof starts from the supersymmetric representation of averages of ratios of characteristic polynomials, which for this profile is a transfer operator on a one-dimensional chain. We show that the transfer chain converges to the heat semigroup of the sigma model, whose noncompact direction produces the Brownian motion in the eigenvector law.

Contents

PART I Introduction and overview 1

1 Introduction 2

2 Overview of the proof 31

PART II Supersymmetry for random band matrices 48

3 Grassmann variables and Berezin integration 50

4 The supersymmetric representation of the band matrix 76

5 The sigma model on a wire 105

6 The sigma model of rank one 136

References 166

PART I

Introduction and overview

SECTION 1

Introduction

1.1 Background

Random band matrices are random Hermitian matrices whose entries vanish, or are negligible, outside a band of width W around the diagonal. They interpolate between Wigner matrices, for W comparable to N, and random Schrödinger operators on a one-dimensional lattice, for W of order one. In physics, they were introduced to describe level statistics intermediate between those of regular and chaotic systems [114] (see [94, Section 5.3]). For band matrices of size N in one dimension, the eigenvectors are expected to be localized on a scale of order W^2, so that one expects a transition at $W \approx N^{1/2}$, with delocalized eigenvectors and the local eigenvalue statistics of the GOE or the GUE, according to the symmetry class, for $W \gg N^{1/2}$, and localized eigenvectors and Poisson statistics for $W \ll N^{1/2}$ [10, (0.2)]. Casati, Molinari, and Izrailev found numerically that the localization of the eigenvectors and the distribution of the level spacings depend on N and on the bandwidth only through the ratio of the square of the bandwidth to N [15, 16]. Fyodorov and Mirlin explained this scaling in [53], where they mapped band matrices with a general profile onto the one-dimensional supersymmetric sigma model of a thick disordered wire, with a localization length proportional to the square of the bandwidth (see [94, (5.39)–(5.41)]). We refer to [10, 128, 129] for surveys.

On the localized side, Schenker proved in [112] that the eigenvectors are localized when $W^8 \ll N$, and the exponent 8 was improved to 7 in [104, Theorem 4]. Localization for $W \ll N^{1/4}$ was proved in [19] for the Wegner orbital model and in [17] for Gaussian band matrices. Recently, Drogin [31] obtained localization at the sharp scale W^2 when $W^2 \ll N$, for real symmetric block band matrices with regular entry distributions. Hislop and Krishna showed that, for real symmetric band matrices with Gaussian entries and $W \sim N^\alpha$ with $\alpha < 1/7$, the limit points of the eigenvalue counts in fixed intervals are Poisson distributed, for almost every energy [67]. Fan and Zou proved convergence of the local eigenvalue statistics to a Poisson process for real symmetric Gaussian block band matrices with $W^2 \log N / N \to 0$ [47, Theorem 1.3].

On the delocalized side, Erdős and Knowles proved quantum diffusion and showed that most eigenvectors have localization length larger than $W^{7/6}$ [40, 41], and delocalization was established for $W \gg N^{4/5}$ in [45]. Bourgade, Yau, and Yin [13] proved quantum unique ergodicity, delocalization, and local GOE statistics for real symmetric

band matrices with $W \geq N^{3/4+\varepsilon}$, using the estimates of Bourgade, Yang, Yau, and Yin [12] and of Yang and Yin [137], after the case $W \asymp N$ [11]; see also [6, 32].

Delocalization down to the threshold was proved recently. Yau and Yin [138] proved delocalization, quantum unique ergodicity, and GUE universality for Gaussian periodic block band matrices with $W \geq N^{1/2+c}$. Erdős and Riabov [44] extended these results to general variance profiles and entry distributions with $W^2 \geq N^{1+c}$, and Fan, Yang, and Yin [48] proved delocalization for general variance profiles with $W \geq N^{1/2+c}$. With supersymmetric methods, M. Shcherbina and T. Shcherbina [119, Theorem 1.1] proved that the two-point correlation function of Gaussian block band matrices converges to that of the GUE when the block size W and the number $n = N/W$ of blocks satisfy $W \geq Cn \log^5 n$. As noted in [31, 47], these results establish the localization–delocalization transition of the eigenvectors and the transition between Poisson and random-matrix local statistics, for large classes of band matrices. None of them describes the critical regime $W^2 \asymp N$, although some of them hold there, as we explain below. Other critical scalings occur at the spectral edge, where the transition to Airy statistics happens at $W \asymp N^{5/6}$ [127] (see [86] for higher dimensions), and the transition at $W \asymp N^{1/2}$ is already visible in the fourth moment of the level density on global scales [100].

In the critical regime $W^2 \asymp N$, the physics prediction is that the statistics of the eigenvalues and eigenvectors are those of the one-dimensional supersymmetric sigma model on a wire of finite length, whose only parameter is the ratio X of the length of the wire to the localization length [38, 53] (see [94, Section 3.2.1]). For the eigenvectors, Fyodorov and Mirlin solved this sigma model exactly for several statistics, including the inverse participation ratio and the distribution of the components [54–56, 93], using the simplification that occurs when the imaginary part of the frequency tends to zero and only the noncompact variable of the sigma model survives [94, Section 3.2.1]. For the eigenvalues, Altland and Fuchs studied the level correlations of a finite wire across the crossover from Wigner–Dyson to Poisson statistics, from numerical solutions of the transfer-matrix equations [4] (see [90, Section V]), and Kettemann studied the analogous autocorrelation of spectral determinants [74, 75]. Closed formulas are known for long wires at all frequencies [87, 90], and, for every ratio of length to localization length, at frequencies that are large compared with the level spacing of a localization volume [90, (25)–(26)]. We compare them with our results in Section 1.3.

The rigorous results that describe the critical regime itself concern characteristic polynomials. In [122, Theorem 1.2], T. Shcherbina considered the band matrix that we study, whose variance profile is the inverse of $\mathrm{Id} - W^2\Delta$ for the Neumann Laplacian Δ on $\{1, \dots, N\}$ (see Definition 1.1 below), with N a fixed multiple of W^2. She proved that the normalized second mixed moment of the characteristic polynomials of this matrix, at two energies at distance of order $1/N$ around E, converges to a matrix element of the semigroup generated by the Laplacian on the sphere S^2 plus a potential. Her parameter

coincides with the parameter $X_N(E) = N/((4-E^2)W^2)$ of our results, since $(2\pi\rho(E))^2 = 4-E^2$ for the semicircle density $\rho(E) = (4-E^2)^{1/2}/(2\pi)$. Her Laplacian, written in the variable $F = 1-2|U_{12}|^2$ of her unitary matrices U, is the operator $\partial_F((1-F^2)\partial_F)$, which reappears below as the compact part of the operator Δ_D of Definition 1.4. This completed the results of [116, 121] on the two sides of the threshold, and analogous critical results were proved for real symmetric block band matrices in [123, Theorem 1.1] and for a non-Hermitian model in [120]. Characteristic polynomials involve only the compact sector of the sigma model, which here is the sphere S^2, while the statistics of eigenvalues and eigenvectors require inverse determinants, and therefore its noncompact sector. The noncompact sector, together with the anticommuting variables, appears in the analysis of the two-point function above the threshold by M. Shcherbina and T. Shcherbina [117, 119].

Several estimates are valid at $W^2 \asymp N$ but do not detect the transition: the density of states of Disertori, Lohmann, and Sodin [29, Theorem 1], the Wegner and Minami estimates of [104], the lower bounds of Erdős and Knowles on the localization length [40, 41], which are uniform in N, their Altshuler–Shklovskii formulas for the density–density correlations on mesoscopic scales $\eta \geq W^{-1/3+c}$ [42, Theorems 2.2–2.3], and, for periodic profiles with sufficiently fast polynomial decay, the delocalization bound $\|\psi\|_\infty^2 \leq N^\tau(W^{-2}+N^{-1})$ of Fan, Yang, and Yin [49, Theorem 2.6], which holds for all bulk eigenvectors ψ and all $W \geq N^c$, for every $\tau > 0$ with probability at least $1-N^{-D}$ for every $D > 0$ and all large N. To our knowledge, no rigorous result on the local eigenvalue process, on the eigenvalue correlation functions on the scale of the level spacing, or on the distribution of the eigenvector mass was previously available at $W^2 \asymp N$.

A rigorous analogue of the critical regime is known for one-dimensional random Schrödinger operators on a box of n sites whose potential has variance of order $1/n$. Kritchevski, Valkó, and Virág [81] showed that their eigenvalues near a bulk energy converge to a point process Sch_τ, where τ depends on the variance of the potential and on the energy. This regime lies between a nearly periodic spectrum, for a potential of smaller variance, and Poisson statistics, for a potential of larger variance [81] (see also [35]). (For CMV matrices with decaying coefficients, Killip and Stoiciu found a transition from Poisson statistics to the clock process [77].) Rifkind and Virág [109, Theorem 1.1] proved that the mass of a uniformly chosen eigenvector of the critical operator, viewed as a measure on $[0,1]$, converges in law to the probability measure with density proportional to $S(\tau(t-U))$, where

$$S(t) = \exp\Big(\frac{\beta'_t}{\sqrt{2}} - \frac{|t|}{4}\Big), \qquad t \in \mathbb{R},$$

β' is a two-sided Brownian motion, and U is uniform on $[0,1]$ and independent of β'. They expect this behavior to be universal for eigenvectors of critical random operators in one dimension, and they state conjectures for several such models [109, Conjectures 1.2–1.3]. Dumaz and Labbé proved one of them, for the continuous Anderson Hamiltonian

on a long interval at energies of the order of its length [34, Theorems 3 and 4], and found the same shape, on an infinite interval, in the localized crossover regime of that operator [35, Theorem 3]; see also [97] for decaying potentials.

In this article, we determine the limit of the eigenvalue process of the band matrix of Definition 1.1 at the critical bandwidth, for energies near zero, together with the macroscopic and local statistics of the eigenvectors. Several of our results confirm predictions obtained by physicists from the sigma model, and our law of the eigenvector mass coincides with the law of Rifkind and Virág; we discuss these relations in Section 1.3.

1.2 Main results

We now state our results. We begin with the definition of the model and the scaling under which we study it, and we then state the results, grouped by the observables they concern.

The model

Definition 1.1. Let $N \geq 2$ be an integer and let $W > 0$. Let Δ denote the Neumann Laplacian on $\{1, \dots, N\}$, which acts on functions $f \colon \{1, \dots, N\} \to \mathbb{C}$ by

$$(\Delta f)(i) = \sum_{j \colon |i-j|=1} \big(f(j) - f(i)\big).$$

The matrix $\mathrm{Id} - W^2\Delta$ is positive definite, since $-\Delta$ is nonnegative. We set $J = (\mathrm{Id} - W^2\Delta)^{-1}$. We let $H = H_{N,W}$ denote the $N \times N$ complex Hermitian random matrix whose entries H_{ij} with $i \leq j$ are independent centered Gaussian random variables, such that H_{ii} is real with variance J_{ii} and, for $i < j$, the real and imaginary parts of H_{ij} are independent, each with variance $J_{ij}/2$. Equivalently,

$$\mathbb{E} H_{ij} H_{kl} = \delta_{il}\delta_{jk} J_{ij}, \qquad i, j, k, l \in \{1, \dots, N\}. \tag{1.1}$$

We call H the *Gaussian band matrix* with bandwidth W, and we call J its *variance profile*.

We let i denote the imaginary unit, since the letter i denotes a site. We use the conventions of Section 1.4 for constants, asymptotic notation, and covariances.

The matrix H belongs to the unitary symmetry class, for which the target space of the sigma model below is simplest. For real symmetric matrices one expects the orthogonal sigma model, with a localization length half as large [38, (52)–(53)] and with the Gaussian factor $\mathbb{E}\gamma^4/(\mathbb{E}\gamma^2)^2 = 3$ of a real Gaussian variable γ in place of the factor 2 in Theorems 1.14 and 1.15 below; the orthogonal analogue of the critical result for characteristic polynomials is in [123].

The profile J is a smoothed version of the usual band structure. We set

$$\theta_W = 2\,\mathrm{arsinh}\Big(\frac{1}{2W}\Big), \qquad J_{\mathbb{Z}}(m) = \frac{e^{-\theta_W|m|}}{\sqrt{1+4W^2}}, \qquad m \in \mathbb{Z}.$$

Then $J_{\mathbb{Z}}(i-j)$ is the kernel of the inverse of $\mathrm{Id} - W^2\Delta_{\mathbb{Z}}$, where $\Delta_{\mathbb{Z}}$ is the Laplacian on $\mathbb{Z}$, and the method of images gives

$$J_{ij} = \sum_{m\in\mathbb{Z}} \big(J_{\mathbb{Z}}(i-j+2mN) + J_{\mathbb{Z}}(i+j-1+2mN)\big), \qquad i,j \in \{1,\dots,N\}. \tag{1.2}$$

All terms in (1.2) are positive, and J decays exponentially in $|i-j|/W$, since $W\theta_W \to 1$ as $W \to \infty$. Since Δ annihilates constants, every row of J sums to one, and $\mathbb{E}\sum_j |H_{ij}|^2 = 1$ for all i, as for a Wigner matrix. When N/W is large, the second moment $\sum_j J_{ij}(i-j)^2$ is of order W^2, uniformly in i, and for the kernel on $\mathbb{Z}$ we have $\sum_{m\in\mathbb{Z}} J_{\mathbb{Z}}(m)m^2 = 2W^2$.

The reason for this choice of profile is that its inverse $\mathrm{Id} - W^2\Delta$ is a nearest-neighbor operator. Our analysis starts from averages of ratios of characteristic polynomials of H, such as (1.4) below. To compute them, one writes each determinant as a Gaussian integral over commuting or anticommuting variables, averages over H, and decouples the resulting quartic interaction by a Hubbard–Stratonovich transformation, which introduces one variable for each site. This variable is a matrix whose entries are commuting or anticommuting variables. We call such a matrix a supermatrix (see Section 3.4). The Gaussian weight of these variables is given by the quadratic form of J^{-1}. Since J^{-1} couples only nearest neighbors, the integrand is a product of factors attached to single sites and to pairs of neighboring sites. The integral is therefore a one-dimensional chain, whose value is a matrix element of a power of a transfer operator; we derive this representation in Proposition 4.11. (This structure underlies earlier supersymmetric work on band matrices with the profile J, such as [27, 29, 115]; see Section 1.3.)

However, the inverse of a sharp profile such as $J_{ij} = (2r_0+1)^{-1}\mathbf{1}_{|i-j|\le r_0}$, when it exists, is not a band matrix, and the same procedure produces an interaction between all pairs of sites. In the physics literature, band matrices with a general normalized profile are mapped onto the same one-dimensional sigma model, in which the profile appears only through its second moment [53] (see [94, (5.39)–(5.41)] and [58, (7)]). More precisely, for a profile whose rows sum to one, we obtain from the lattice sigma model of [58, (7)], as in Section 1.3, the parameter $2N/((4-E^2)\sum_r J(r)r^2)$, where the second moment is taken away from the ends of the chain. Our profile has $\sum_r J(r)r^2 = 2W^2$ there, and this parameter becomes $N/((4-E^2)W^2) = X_N(E)$, with $X_N(E)$ as in (1.3) below. The sharp profile of half-width r_0 has second moment $r_0(r_0+1)/3$, and Figure 1.1 compares the two profiles when their second moments agree. Our results should therefore describe general profiles, and we conjecture that they hold for the sharp profile of half-width r_0, with W^2 replaced by $r_0(r_0+1)/6$ in $X_N(E)$. Our proof, however, uses the nearest-neighbor

structure of J^{-1} in an essential way. The Neumann boundary condition corresponds to a wire with two ends [94, Section 3.2.1].

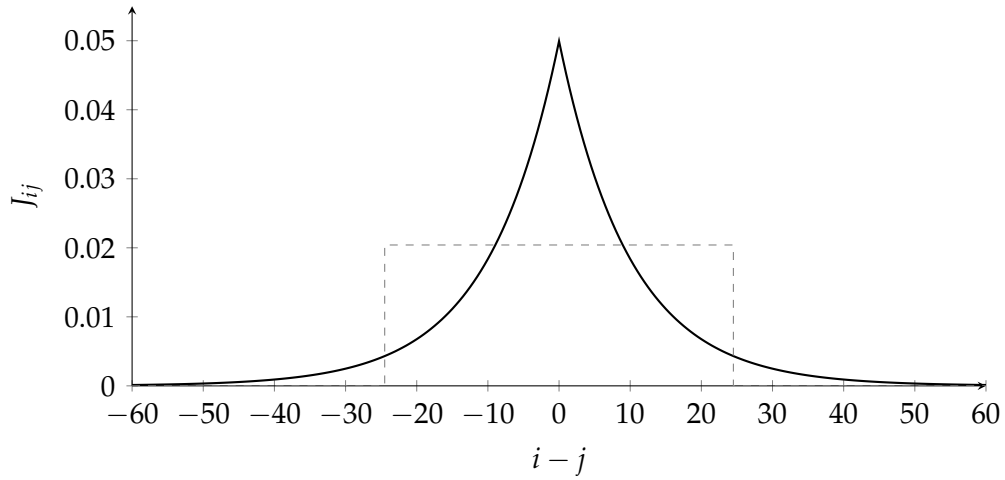


Figure 1.1. The kernel $J_{\mathbb{Z}}(i-j)$ of (1.2) for $W = 10$ (solid line), which is a row of the variance profile J far from the ends of the chain, up to exponentially small terms, and the sharp profile $(2r_0+1)^{-1}\mathbf{1}_{|i-j|\le r_0}$ with $r_0 = 24$ (dashed gray line). Both have total mass one and the second moment $2W^2 = r_0(r_0+1)/3 = 200$.

The law of H has a positive density on the space of Hermitian matrices, since the entries of J are positive. In particular, the eigenvalues of H are simple almost surely. We let $\lambda_1 < \cdots < \lambda_N$ denote them, and we let $u_1, \dots, u_N$ denote corresponding normalized eigenvectors, with coordinates $u_\alpha(i)$. We set $w_\alpha(i) = |u_\alpha(i)|^2$. These weights do not depend on the choice of phases, and $\sum_i w_\alpha(i) = 1$ and $\sum_\alpha w_\alpha(i) = 1$. Throughout, α denotes an eigenvalue index and i, j denote sites. For $E \in (-2, 2)$, we set

$$\begin{aligned} s_E &= \sqrt{1 - E^2/4}, \qquad \rho(E) = \frac{s_E}{\pi}, \qquad a_N = N\rho(E), \\ X_N(E) &= \frac{N}{(4 - E^2)W^2}, \qquad x_\alpha = a_N(\lambda_\alpha - E). \end{aligned} \tag{1.3}$$

By Lemma 7.3, the expected density of states of H converges to the semicircle density $\rho(E) = (4 - E^2)^{1/2}/(2\pi)$, and the unfolded eigenvalues x_α have mean spacing one near zero. The parameter $X_N(E)$ compares N with W^2. In the physics literature, W^2 is the order of the localization length of the eigenvectors when $N = \infty$ [15, 53], and $X_N(E) = N/\xi$ is the length of the chain in units of the localization length $\xi = (4 - E^2)W^2$ (see Section 1.3). For $N \ll W^2$ one expects the local statistics of the GUE and delocalized eigenvectors, and for $N \gg W^2$ one expects Poisson statistics and localized eigenvectors (see Section 1.1). This paper concerns the critical regime $N \asymp W^2$, in which $X_N(E)$ stays bounded and bounded away from zero, and in which features of the two regimes appear on different scales.

The paper has four parts, which are described in Section 1.3. Part I consists of this section and the overview of the proof, Part II introduces supersymmetry and the sigma model, Part III derives the results below from the sigma model and from the interface theorem, Theorem 7.6, and Part IV proves the interface theorem. This theorem states that averages of ratios of characteristic polynomials of H, and their derivatives in perturbations at single sites, converge to quantities of the sigma model, and that the diagonal entries of the resolvent of H have bounded moments and, at sites away from 1 and N, the mean that the semicircle law predicts; we describe it in Section 2.1. Every result of Section 1.2 uses it.

We now fix the energies and the limits that our results concern. Theorem 7.6 involves two universal energies $0 < E_* < E_{\text{out}} \leq 1/2$, the inner and the outer energy bound of Remark 7.7. Our analysis of the transfer operator in Part IV applies for $|E| \leq E_{\text{out}}$, and this bound will be fixed in (20.2). The inner bound E_* will be fixed in Section 20.1. The margin $E_{\text{out}} - E_*$ absorbs the shifts of the energy by $O(1/N)$ in the proof of Theorem 7.6, and Part III uses it to control the eigenvalues far from E. Both bounds are the same for all our results, and in particular they do not depend on the number of characteristic polynomials in the averages below or on the order of any moment. Their values are not explicit, since they depend on universal constants in Part IV that we do not compute. All our results are stated for $|E| \leq E_*$. These restrictions come from the estimates of Part IV, and we expect the results to hold for all $|E| < 2$.

Definition 1.2. Let $X \in (0, \infty)$. We say that (N, W, E) tends to the *critical limit at* X if $N, W \to \infty$ and $E = E_{N,W}$ satisfies $|E| \leq E_*$ and $X_N(E) \to X$. We say that a statement holds *uniformly in the critical regime* if for all $0 < X_- \leq X_+ < \infty$ it holds uniformly over all (N, W, E) with $|E| \leq E_*$ and $X_- \leq X_N(E) \leq X_+$, as $N \to \infty$.

In these limits the energy E may depend on N and W. Several statements below concern limits $X \downarrow 0$, $X \to \infty$, or $L \to \infty$ of objects defined below for fixed $X \in (0, \infty)$, such as the limiting point process Ξ_X and the kernels K_X, which arise in the critical limit at X. In all such statements the critical limit is taken first, and the second limit is a statement about the limiting objects. We do not consider joint limits in which $X_N(E) \to 0$ or $X_N(E) \to \infty$ as $N \to \infty$.

The eigenvalue process

We identify a locally finite point configuration on $\mathbb{R}$ with its counting measure. We equip the set of locally finite counting measures on $\mathbb{R}$ with the vague topology, the coarsest topology for which the maps $\zeta \mapsto \int f \, \mathrm{d}\zeta$ are continuous for all continuous functions f with compact support. We consider the unfolded eigenvalue process

$$\Xi_N = \sum_{\alpha=1}^{N} \delta_{x_\alpha}.$$

Our description of the limit of Ξ_N is based on averages of ratios of characteristic polynomials. Fix $\varepsilon > 0$ and an integer $k \geq 1$. For $\mathbf{u} = (u_1, \dots, u_k) \in \mathbb{R}^k$ with $\sum_j u_j = 0$ and $\mathbf{b} = (b_1, \dots, b_k) \in (0, \infty)^k$, we set

$$T_N^{(k)}(\mathbf{u}, \mathbf{b}) = \mathbb{E} \prod_{\alpha=1}^{N} \prod_{j=1}^{k} \frac{(x_\alpha - u_j)^2 + b_j^2}{x_\alpha^2 + \varepsilon^2}. \tag{1.4}$$

We call $T_N^{(k)}$ the *source ratio*, $(\mathbf{u}, \mathbf{b})$ the *sources*, k the *source rank*, and ε the *damping*. The dependence of $T_N^{(k)}$ on ε is not displayed. Writing $z_j = E + (u_j + \mathrm{i}b_j)/a_N$ and $z_0 = E + \mathrm{i}\varepsilon/a_N$, we have

$$T_N^{(k)}(\mathbf{u}, \mathbf{b}) = \mathbb{E} \frac{\prod_{j=1}^{k} \det(H - z_j) \det(H - \bar{z}_j)}{\big(\det(H - z_0) \det(H - \bar{z}_0)\big)^k}.$$

Each determinant in the numerator is a Gaussian integral over anticommuting (Grassmann) variables, and each determinant in the denominator is a Gaussian integral over complex variables (see Section 3.1). Averages of this kind are the quantities that the supersymmetric method computes. Since $\sum_j u_j = 0$, we have

$$\prod_{j=1}^{k} \frac{(x - u_j)^2 + b_j^2}{x^2 + \varepsilon^2} = 1 + O(x^{-2}), \qquad |x| \to \infty,$$

and the corresponding product over the points of a process of intensity one converges absolutely.

The limits of $T_N^{(k)}$ are partition functions of the supersymmetric sigma model of a disordered wire [38, 39], which Fyodorov and Mirlin derived for band matrices [53].

Definition 1.3. Let $X > 0$, $\varepsilon > 0$, $k \geq 1$, $\mathbf{u} \in \mathbb{R}^k$ with $\sum_j u_j = 0$, and $\mathbf{b} \in (0, \infty)^k$. We let $\mathcal{Z}_{X,\varepsilon}^{(k)}(\mathbf{u}, \mathbf{b})$ denote the partition function of the supersymmetric sigma model on $[0, 1]$ with coupling X and sources $(\mathbf{u}, \mathbf{b})$, which is defined in Definition 5.16.

The reader may treat $\mathcal{Z}_{X,\varepsilon}^{(k)}$ as an explicit superintegral, defined in Section 5, and Theorem 1.6 below can be read without its definition.

We describe it briefly. We call a $4k \times 4k$ matrix a supermatrix of size $(2k|2k)$ if its two diagonal blocks of size $2k$, the BB and the FF block, have commuting entries and its two off-diagonal blocks have anticommuting entries (see Section 3.4). We let $\operatorname{diag}(A \mid D)$ denote the supermatrix with BB block A, FF block D, and vanishing off-diagonal blocks, and we let STr denote the supertrace, given by $\operatorname{STr} M = \operatorname{Tr} A - \operatorname{Tr} D$ for a supermatrix M with BB block A and FF block D. Let $\mathcal{M}_k$ denote the symmetric superspace $U(k, k|2k)/(U(k|k) \times U(k|k))$, a quotient of Lie supergroups that Section 5 realizes as

a set of supermatrices Q of size $(2k|2k)$ with $Q^2 = \mathrm{Id}$. Its underlying manifold is the product of the noncompact symmetric space $U(k,k)/(U(k)\times U(k))$ and the compact symmetric space $U(2k)/(U(k)\times U(k))$. We call the functions on $\mathcal{M}_k$ superfunctions. Besides a point of the underlying manifold, they depend on $4k^2$ anticommuting variables (see Section 3). The partition function $\mathcal{Z}^{(k)}_{X,\varepsilon}(\mathbf{u},\mathbf{b})$ is the Berezin integral over $\mathcal{M}_k$, an integral over the commuting and the anticommuting variables, of the heat semigroup at time one, applied to the constant function one, of the operator $X\mathsf{C}_k$ plus a potential. Here the Casimir operator C_k is minus the invariant Laplace–Beltrami operator of $\mathcal{M}_k$, and the potential is linear in Q and depends on ε, $\mathbf{u}$, and $\mathbf{b}$.

Precisely, let $\Lambda_{\mathrm{B}} = \mathrm{diag}(\mathrm{Id}_k, -\mathrm{Id}_k)$, and let $\Gamma = \mathrm{diag}(\varepsilon\Lambda_{\mathrm{B}} \,|\, \Gamma_{\mathrm{F}})$ denote the diagonal supermatrix with

$$\Gamma_{\mathrm{F}} = \mathrm{diag}(b_1 - \mathrm{i}u_1, \dots, b_k - \mathrm{i}u_k, -b_1 - \mathrm{i}u_1, \dots, -b_k - \mathrm{i}u_k).$$

The diagonal entries of $E + \mathrm{i}\Gamma/a_N$ are the spectral parameters $z_0, \bar{z}_0$ and $z_j, \bar{z}_j$ of $T^{(k)}_N$. Then

$$\mathcal{Z}^{(k)}_{X,\varepsilon}(\mathbf{u},\mathbf{b}) = \int e^{-(X\mathsf{C}_k + \pi\,\mathrm{STr}(Q\Gamma))}1,$$

where $\int$ is the invariant Berezin integral over $\mathcal{M}_k$ of Definition 5.6, normalized so that $\int e^{-d_k} = 1$ with $d_k(Q) = \frac{1}{2}\mathrm{STr}(\Lambda Q)$ and $\Lambda = \mathrm{diag}(\Lambda_{\mathrm{B}} \,|\, \Lambda_{\mathrm{B}})$. The time variable of the semigroup is the position along the chain, rescaled to $[0,1]$. We will use two properties of $\mathcal{Z}^{(k)}_{X,\varepsilon}$ in this section. By Corollary 5.20, we have $\mathcal{Z}^{(k)}_{X,\varepsilon}(0,\varepsilon\mathbf{1}) = 1$, in agreement with $T^{(k)}_N(0,\varepsilon\mathbf{1}) = 1$. At rank one, the second derivative of $\mathcal{Z}^{(1)}_{X,\varepsilon}(0,b)$ in b at $b = \varepsilon$ is given by the function I_X of Definition 1.4 below.

The underlying manifold of $\mathcal{M}_1$ is the product of the hyperbolic plane and the sphere S^2. The rank-one computations of Section 6 reduce to functions of two coordinates, the hyperbolic cosine $B \geq 1$ of the distance to a base point of the hyperbolic plane and the cosine $F \in [-1,1]$ of the distance to a base point of S^2. The operator Δ_D in the next definition is the Laplace–Beltrami operator of the product acting on functions of (B,F).

Definition 1.4. Let $D = [1,\infty)\times[-1,1]$, with coordinates (B,F), and set $d = B - F$. We define the differential operator

$$\Delta_D = \partial_B\big((B^2-1)\partial_B\big) + \partial_F\big((1-F^2)\partial_F\big).$$

Let $X > 0$. We let $(\tilde{B}_t, \tilde{F}_t)_{t\geq 0}$ denote the diffusion on D with generator $X\Delta_D$, and we let $\mathbb{E}_p$ denote the expectation when $(\tilde{B}_0, \tilde{F}_0) = p \in D$. For $z \in \mathbb{C}$ with $\mathrm{Re}\, z > 0$, $t \geq 0$, and bounded measurable functions f on D, we set

$$\big(e^{t(X\Delta_D - zd)}f\big)(p) = \mathbb{E}_p\Big[f(\tilde{B}_t,\tilde{F}_t)\exp\Big(-z\int_0^t(\tilde{B}_s - \tilde{F}_s)\,\mathrm{d}s\Big)\Big],$$

and we define the *singlet state* $\Psi_t^{X,z}\colon D \to \mathbb{C}$ by

$$\Psi_t^{X,z} = 1 - zd \int_0^t e^{v(X\Delta_D - zd)} 1 \, \mathrm{d}v. \tag{1.5}$$

For $0 < x \le y < 1$ and $\operatorname{Re} z > 0$, we define

$$K_X(x,y;z) = \int_D \Psi_x^{X,z} \left(e^{(y-x)(X\Delta_D - zd)} \Psi_{1-y}^{X,z}\right) \mathrm{d}B \, \mathrm{d}F, \tag{1.6}$$

$$\widetilde{K}_X(x;z) = \int_D \frac{B}{d} \, \Psi_x^{X,z} \Psi_{1-x}^{X,z} \, \mathrm{d}B \, \mathrm{d}F, \tag{1.7}$$

$$I_X(z) = \int_0^1 \int_x^1 K_X(x,y;z) \, \mathrm{d}y \, \mathrm{d}x, \tag{1.8}$$

and we set $K_X(x,y;z) = K_X(y,x;z)$ for $0 < y < x < 1$. We call $\widetilde{K}_X$ the *contact kernel*.

The coordinates $\tilde{B}_t$ and $\tilde{F}_t$ are independent one-dimensional diffusions, with generators $X\partial_B((B^2-1)\partial_B)$ on $[1,\infty)$ and $X\partial_F((1-F^2)\partial_F)$ on $[-1,1]$. A computation of their scale functions and speed measures shows that their finite endpoints are entrance boundaries in Feller's classification. This implies that the diffusion is determined by its generator and does not reach the boundary of D at positive times. The integrals in (1.6)–(1.8) converge absolutely, and K_X, $\widetilde{K}_X$, and I_X are holomorphic in the half-plane $\operatorname{Re} z > 0$ (see Sections 6 and 10). The singularity of B/d in (1.7) is at the corner $(1,1)$, where it is integrable.

Remark 1.5. The time variables in Definition 1.4 are positions along the chain. In (1.6), the diffusion runs for time x from the left end of $[0,1]$, for time $y-x$ between the two points x and y, and for time $1-y$ from the right end. The singlet state has the following probabilistic meaning. A direct computation shows that $1/d$ is harmonic for Δ_D away from the corner $(1,1)$. Let $(B_t, F_t)_{t\ge 0}$ denote the Doob transform of $(\tilde{B}_t, \tilde{F}_t)$ by $1/d$, which is the diffusion with generator

$$Xd\,\Delta_D\, d^{-1} = X\Big(\Delta_D - \frac{2(B^2-1)}{d}\partial_B + \frac{2(1-F^2)}{d}\partial_F\Big),$$

stopped when it reaches the corner. We call it the singlet diffusion. For $\operatorname{Re} z > 0$, the defining property of the Doob transform and (1.5) give

$$\Psi_t^{X,z}(p) = \mathbb{E}_p^{\mathrm{s}}\Big[\exp\Big(-z\int_0^t (B_v - F_v)\,\mathrm{d}v\Big)\Big],$$

where $\mathbb{E}_p^{\mathrm{s}}$ denotes the expectation when $(B_0,F_0) = p$ (see Lemma 6.5). Consequently $|\Psi_t^{X,z}| \le \Psi_t^{X,\operatorname{Re} z} \le 1$ and $\Psi_t^{X,z}(1,1) = 1$, since d vanishes at the corner. By Lemma 5.19, the Berezin integral over $\mathcal{M}_1$ of a superfunction that is annihilated by the odd vector

field (supercharge) of that lemma equals the value of the superfunction at the point Λ of $\mathcal{M}_1$, and this point corresponds to the corner. We explain in Section 6 how $\Psi_t^{X,z}$ and the kernels arise from the sigma model. In the notation of the physics literature, we have $X\Delta_D - zd = -2Xd^{-1}\mathcal{H}d$, where $\mathcal{H}$ is the transfer-matrix Hamiltonian of the unitary wire [71, (12)], $d^{-1}\mathcal{H}d$ is the operator of [71, (18)–(20)], and their dimensionless frequency is set to $z/(2X)$; time there is position in units of the localization length.

Our first main result is the following theorem, which identifies the limit of Ξ_N.

Theorem 1.6. *Let $X \in (0,\infty)$. In the critical limit at X, the following statements hold.*

(a) The point processes Ξ_N converge in law, in the vague topology, to a point process Ξ_X on $\mathbb{R}$ that is simple, stationary, invariant under the reflection $x \mapsto -x$, and of intensity one. The law of Ξ_X depends only on X.

(b) For all $\varepsilon > 0$, $k \geq 1$, $\mathbf{u} \in \mathbb{R}^k$ with $\sum_j u_j = 0$, and $\mathbf{b} \in (0,\infty)^k$, we have $T_N^{(k)}(\mathbf{u},\mathbf{b}) \to \mathcal{Z}_{X,\varepsilon}^{(k)}(\mathbf{u},\mathbf{b})$ and

$$\mathbb{E} \prod_{x\in\Xi_X} \prod_{j=1}^{k} \frac{(x-u_j)^2 + b_j^2}{x^2+\varepsilon^2} = \mathcal{Z}_{X,\varepsilon}^{(k)}(\mathbf{u},\mathbf{b}). \tag{1.9}$$

For every fixed $\varepsilon > 0$, the values of the left side of (1.9) for all k, $\mathbf{u}$, and $\mathbf{b}$ as above determine the law of Ξ_X.

(c) For all $\varepsilon > 0$,

$$\partial_t T_N^{(1)}(0,\varepsilon+t)\big|_{t=0} \longrightarrow 2\pi, \qquad \partial_t^2 T_N^{(1)}(0,\varepsilon+t)\big|_{t=0} \longrightarrow 4\pi^2\big(1 + I_X(2\pi\varepsilon)\big). \tag{1.10}$$

The proof of Theorem 1.6 is given in Section 8. The matrices appear in the proof of part (b) only through Theorem 7.6, which is the main result of Part IV. We derive parts (a) and (c) from part (b), from the counting estimates of Lemma 7.2, which follow from known results, and, for part (c), from a computation in the sigma model of rank one. By part (b), a single damping ε suffices to determine the law. The reason is that the logarithms of the rational functions in (1.9), which all have the same poles $\pm i\varepsilon$, approximate every nonnegative continuous function f with compact support in the norm $\sup_x (x^2+\varepsilon^2)|f(x)|$. After the change of variables $x = \varepsilon\tan(\vartheta/2)$, this follows from Fejér's theorem.

The limits in part (c) concern sources of rank one. Let $P_\varepsilon(v) = \varepsilon/(\pi(v^2+\varepsilon^2))$ denote the Poisson kernel. Differentiating the product in (1.4), we find

$$\partial_t T_N^{(1)}(0,\varepsilon+t)\big|_{t=0} = 2\pi\mathbb{E}\sum_{\alpha=1}^{N} P_\varepsilon(x_\alpha),$$

and the first limit in (1.10) expresses that Ξ_N has asymptotic intensity one. The second derivative is a two-point function of Ξ_N smoothed at scale ε. Together with stationarity, it identifies the pair correlation of Ξ_X, as follows. Theorem 1.6 confirms the prediction,

recalled in Section 1.1, that the local eigenvalue statistics at the critical bandwidth form a family of laws indexed by X. We call a positive Radon measure ν on $\mathbb{R}$ translation-bounded if $\sup_{s\in\mathbb{R}} \nu([s,s+1]) < \infty$. We call the measure ν_X of the next corollary the *pair measure* of Ξ_X. It is the reduced second factorial moment measure of Ξ_X in the sense of [22, Section 8.1].

Corollary 1.7. *Let $X \in (0,\infty)$. There exists an even, positive, and translation-bounded Radon measure ν_X on $\mathbb{R}$ such that*

$$\mathbb{E} \sum_{\substack{x,y\in\Xi_X \\ x\neq y}} f(x,y) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(v,v+s)\, \nu_X(\,\mathrm{d}s)\,\mathrm{d}v \tag{1.11}$$

for all nonnegative measurable functions f on $\mathbb{R}^2$. It satisfies

$$(P_\varepsilon * \nu_X)(v) = 1 + \operatorname{Re} I_X(\pi\varepsilon - \mathrm{i}\pi v) - P_\varepsilon(v), \qquad \varepsilon > 0,\ v \in \mathbb{R}, \tag{1.12}$$

and it is the unique translation-bounded measure with this property. Further, for all bounded intervals $A_1, A_2 \subset \mathbb{R}$, in the critical limit at X,

$$\operatorname{Cov}\big(\Xi_N(A_1), \Xi_N(A_2)\big) \longrightarrow |A_1 \cap A_2| + \int_{\mathbb{R}} |A_1 \cap (A_2 - s)|\, \nu_X(\,\mathrm{d}s) - |A_1||A_2|, \tag{1.13}$$

and the right side equals $\operatorname{Cov}(\Xi_X(A_1), \Xi_X(A_2))$.

Corollary 1.7 is proved in Section 14. By (1.12), the function $1 + \operatorname{Re} I_X(\pi\varepsilon - \mathrm{i}\pi v)$ is the Poisson smoothing at scale ε of the two-point function $\delta_0 + \nu_X$ of Ξ_X, including its diagonal term. For $L \geq 0$, we set

$$\Sigma^2_X(L) = L + \int_{-L}^{L} (L - |s|)\, \nu_X(\,\mathrm{d}s) - L^2. \tag{1.14}$$

Applying Corollary 1.7 with $A_1 = A_2 = [c, c+L]$, we see that

$$\Sigma^2_X(L) = \operatorname{Var} \Xi_X([c,c+L]) = \lim \operatorname{Var} \Xi_N([c,c+L]), \qquad c \in \mathbb{R},$$

where the limit is the critical limit at X. We call $\Sigma^2_X(L)$ the *number variance* of Ξ_X.

Eigenvectors

We select the eigenvectors of H through their eigenvalues, with a smooth window. We fix a nonnegative function $\phi \in C_c^\infty((-1,1))$ with $\int \phi = 1$ for the rest of the paper. For $E \in \mathbb{R}$ and $r > 0$, we set $\phi_{E,r}(\lambda) = r^{-1}\phi((\lambda - E)/r)$.

Definition 1.8. Let $E \in (-2,2)$ and $r > 0$, and let $w_\alpha = (w_\alpha(1), \dots, w_\alpha(N))$. For each α, let $\Phi_\alpha = \Phi(\lambda_\alpha, w_\alpha)$, where $\Phi\colon \mathbb{R} \times [0,1]^N \to \mathbb{R}$ is a measurable function that is bounded or nonnegative. We define the *annealed selection* of Φ_α at energy E and scale r by

$$\langle \Phi_\alpha \rangle_{E,r} = \frac{\mathbb{E} \sum_\alpha \phi_{E,r}(\lambda_\alpha) \Phi_\alpha}{\mathbb{E} \sum_\alpha \phi_{E,r}(\lambda_\alpha)},$$

and, for $i \in \{1, \dots, N\}$, its *site-rooted* version by

$$\langle \Phi_\alpha \rangle_{E,r}^i = \frac{\mathbb{E} \sum_\alpha \phi_{E,r}(\lambda_\alpha) w_\alpha(i) \Phi_\alpha}{\mathbb{E} \sum_\alpha \phi_{E,r}(\lambda_\alpha) w_\alpha(i)}.$$

The denominators are positive, since the law of H has a positive density. The two selections are expectations under probability measures on pairs (H, α). The first tilts the law of H by $\sum_\alpha \phi_{E,r}(\lambda_\alpha)$, and given H it chooses α with probability proportional to $\phi_{E,r}(\lambda_\alpha)$. The second tilts further by $w_\alpha(i)$. They are ratios of expectations, not expectations of random ratios (this is the sense of the word annealed), and they are well defined when r is much smaller than the spacing of the eigenvalues near E, in which case a typical sample of H has no eigenvalue in the support of $\phi_{E,r}$. Our results hold uniformly for $0 < r \le \bar{r}_N$, for every deterministic sequence $\bar{r}_N \to 0$. This covers all windows that are small on the macroscopic scale, including windows much smaller than the spacing.

We let $\mathcal{P}([0,1])$ denote the space of Borel probability measures on $[0,1]$, with the topology of weak convergence. We set $\mu(f) = \int f \, \mathrm{d}\mu$ for $\mu \in \mathcal{P}([0,1])$ and bounded measurable functions f on $[0,1]$. We define the *mass measure* of u_α by

$$\mu_\alpha^{(N)} = \sum_{i=1}^N w_\alpha(i) \delta_{i/N} \in \mathcal{P}([0,1]).$$

Its limit is described by the following Brownian objects.

Definition 1.9. Let $X > 0$, and let $\beta = (\beta_t)_{t \in [0,1]}$ be a standard Brownian motion, with law $\mathbb{P}$ and expectation $\mathbb{E}$. We set

$$g_X(t) = \sqrt{2X}\beta_t - Xt, \qquad Z_X = \int_0^1 e^{g_X(t)} \, \mathrm{d}t, \qquad \mu_X(\mathrm{d}t) = \frac{e^{g_X(t)}}{Z_X} \, \mathrm{d}t.$$

Since $\mathbb{E} e^{g_X(t)} = 1$ for all t, we have $\mathbb{E} Z_X = 1$. We let $\mathbb{Q}_X$ denote the probability measure defined by $\mathrm{d}\mathbb{Q}_X = Z_X \, \mathrm{d}\mathbb{P}$. For $k \ge 1$ and $x_1, \dots, x_k \in [0,1]$, we set

$$M_k^X(x_1, \dots, x_k) = \mathbb{E}\left[\frac{\exp\big(g_X(x_1) + \cdots + g_X(x_k)\big)}{Z_X^{k-1}} \right], \qquad \mathcal{C}_X(x_1, x_2) = M_2^X(x_1, x_2).$$

The random measure μ_X is the normalized exponential of a Brownian motion with drift, and $\mathbb{Q}_X$ tilts its law by the total mass Z_X. Using the definition of $\mathbb{Q}_X$, we see that M_k^X is the density of the k-th moment measure of μ_X under $\mathbb{Q}_X$,

$$\mathbb{E}_{\mathbb{Q}_X} \int_{[0,1]^k} f \, \mathrm{d}\mu_X^{\otimes k} = \int_{[0,1]^k} f(x) M_k^X(x) \, \mathrm{d}x$$

for all nonnegative measurable functions f on $[0,1]^k$. In particular $M_1^X = 1$. We call $\mathcal{C}_X$ the *eigenvector correlation*.

Theorem 1.10. *Let $\bar{r}_N \to 0$ be a deterministic sequence, and let $\Phi\colon \mathcal{P}([0,1]) \to \mathbb{R}$ be continuous. Then, uniformly in the critical regime and in $0 < r \le \bar{r}_N$,*

$$\big\langle \Phi(\mu_\alpha^{(N)}) \big\rangle_{E,r} - \mathbb{E}_{\mathbb{Q}_{X_N(E)}} \big[\Phi(\mu_{X_N(E)}) \big] \longrightarrow 0.$$

In particular, in the critical limit at X, the law of $\mu_\alpha^{(N)}$ under $\langle \cdot \rangle_{E,r}$ converges weakly to the law of μ_X under $\mathbb{Q}_X$, for every choice of $r = r_N \in (0, \bar{r}_N]$.

Theorem 1.10 is proved in Section 12. It describes the distribution of the mass of an eigenvector, selected by its eigenvalue, on the macroscopic scale. In the critical regime the mass is spread over the whole chain, but it is not flat. Its density is the exponential of a Brownian motion with drift, normalized to have total mass one, under the law tilted by the total mass. This confirms the law of the envelope of an eigenfunction derived by Fyodorov and Mirlin [56], in the form of [94, (3.32)–(3.34)]. The same law, as a law of exponential functionals of Brownian motion, was proved by Rifkind and Virág for the critical one-dimensional random Schrödinger operator recalled in Section 1.1, with their parameter τ equal to $4X$ [109, Theorem 1.1]; we check this identification in Remark 12.7. Theorem 1.10 shows that band matrices at the critical bandwidth have the same coarse eigenvector law. The Brownian motion comes from the noncompact direction B of the sigma model. In Section 10, we show that the diffusion of Definition 1.4 becomes a geometric Brownian motion on the scale where B is large, with logarithm given by the process g_X. As $X \downarrow 0$, the measure μ_X tends to the Lebesgue measure, which is the macroscopic profile of the eigenvectors of the GUE. As $X \to \infty$, the mass concentrates near a uniformly distributed point, as expected for localized eigenvectors (see Proposition 1.16). The coarse law does not control the mass at the scale of single sites, since functions such as $\sum_i w_\alpha(i)^2$ are not continuous functions of $\mu_\alpha^{(N)}$. We treat these in Theorem 1.15.

The next corollary concerns the mean and the covariance of the mass. We let $\mathrm{Cov}_{\mathbb{Q}_X}$ denote the covariance under $\mathbb{Q}_X$, and for observables Φ_α and Φ'_α as in Definition 1.8, we define their covariance under the annealed selection by

$$\mathrm{Cov}_{E,r}(\Phi_\alpha, \Phi'_\alpha) = \langle \Phi_\alpha \Phi'_\alpha \rangle_{E,r} - \langle \Phi_\alpha \rangle_{E,r} \langle \Phi'_\alpha \rangle_{E,r}.$$

Corollary 1.11. *Let $X > 0$, and let $f, g\colon [0,1] \to \mathbb{R}$ be bounded and Riemann integrable. Then*

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}_X} \mu_X(f) &= \int_0^1 f(t)\,\mathrm{d}t, \\ \mathrm{Cov}_{\mathbb{Q}_X}\big(\mu_X(f), \mu_X(g)\big) &= \int_0^1 \int_0^1 f(x) g(y) \big(\mathcal{C}_X(x,y) - 1\big)\,\mathrm{d}x\,\mathrm{d}y. \end{aligned} \tag{1.15}$$

Further, let $\bar{r}_N \to 0$ be a deterministic sequence. In the critical limit at X, uniformly in $0 < r \le \bar{r}_N$, the quantities $\langle \mu_\alpha^{(N)}(f) \rangle_{E,r}$ and $\mathrm{Cov}_{E,r}(\mu_\alpha^{(N)}(f), \mu_\alpha^{(N)}(g))$ converge to the corresponding quantities in (1.15)*.*

The exact identities in (1.15) follow from Definition 1.9 and Tonelli's theorem, and the matrix limits are proved in Section 12. In particular $\mathbb{E}_{\mathbb{Q}_X} \mu_X$ is the Lebesgue measure on $[0,1]$, and a selected eigenvector is flat on average. The function $\mathcal{C}_X - 1$ is the covariance density of the mass.

We now describe the mass as seen from a point of the chain, which we call the root. There are two natural choices of the root, at random according to the mass of the eigenvector, or at a given site under the site-rooted selection. For the first, we let $\overline{\mathbb{Q}}_X$ denote the probability measure on pairs (β, U), where β is a path on $[0,1]$ and $U \in [0,1]$, defined by

$$\overline{\mathbb{Q}}_X(\,\mathrm{d}\beta,\,\mathrm{d}U) = e^{g_X(U)}\,\mathbb{P}(\,\mathrm{d}\beta)\,\mathrm{d}U.$$

Since $\int_0^1 e^{g_X(U)}\,\mathrm{d}U = Z_X$ and $\mathbb{E}e^{g_X(U)} = 1$, the measure $\overline{\mathbb{Q}}_X$ has first marginal $\mathbb{Q}_X$ and a uniform second marginal. For the second, we let $x \in (0,1)$ and set

$$p_{X,x}(t) = \frac{\exp\big(\sqrt{2X}(\beta_t - \beta_x) - X|t-x|\big)}{\int_0^1 \exp\big(\sqrt{2X}(\beta_s - \beta_x) - X|s-x|\big)\,\mathrm{d}s}, \qquad t \in [0,1]. \tag{1.16}$$

For the large-X limit, we let $(\beta_s)_{s\in\mathbb{R}}$ denote a two-sided Brownian motion, so that $(\beta_s)_{s\ge 0}$ and $(\beta_{-s})_{s\ge 0}$ are independent standard Brownian motions, and we set

$$p_\infty(s) = \frac{e^{\sqrt{2}\beta_s - |s|}}{\int_{\mathbb{R}} e^{\sqrt{2}\beta_{s'} - |s'|}\,\mathrm{d}s'}, \qquad s \in \mathbb{R}. \tag{1.17}$$

The integral in the denominator is finite almost surely, since $\beta_s/s \to 0$ as $|s| \to \infty$. For $X > 0$ and $t_0 \in [0,1]$, we let $\sigma_{X,t_0}(t) = X(t - t_0)$. We let $\mathcal{P}(\mathbb{R})$ denote the space of Borel probability measures on $\mathbb{R}$, with the topology of weak convergence.

Theorem 1.12. *Let $X \in (0,\infty)$, and let $\bar{r}_N \to 0$ be a deterministic sequence. The following statements hold.*

(a) In the critical limit at X, uniformly in $0 < r \le \bar{r}_N$, for every continuous function $\Phi\colon \mathcal{P}([0,1]) \times [0,1] \to \mathbb{R}$,

$$\Big\langle \int_0^1 \Phi\big(\mu_\alpha^{(N)}, t\big)\,\mu_\alpha^{(N)}(\,\mathrm{d}t) \Big\rangle_{E,r} \longrightarrow \mathbb{E}_{\overline{\mathbb{Q}}_X}\big[\Phi(\mu_X, U)\big].$$

(b) Let $i = i_N$ satisfy $i/N \to x \in (0,1)$. In the critical limit at X, uniformly in $0 < r \le \bar{r}_N$, for every continuous function $\Phi\colon \mathcal{P}([0,1]) \to \mathbb{R}$,

$$\big\langle \Phi(\mu_\alpha^{(N)}) \big\rangle_{E,r}^{i} \longrightarrow \mathbb{E}\big[\Phi\big(p_{X,x}(t)\,\mathrm{d}t\big)\big].$$

(c) As $X \to \infty$, the law of $\big((\sigma_{X,U})_*\mu_X, U\big)$ under $\overline{\mathbb{Q}}_X$ converges weakly to the law of $(p_\infty(s)\,\mathrm{d}s, U')$, where U' is uniform on $[0,1]$ and independent of β. For every $x \in (0,1)$, the law of $(\sigma_{X,x})_*(p_{X,x}(t)\,\mathrm{d}t)$ under $\mathbb{P}$ converges weakly to the law of $p_\infty(s)\,\mathrm{d}s$, as random elements of $\mathcal{P}(\mathbb{R})$.

Theorem 1.12 is proved in Section 12. In part (a), the integral against $\mu_\alpha^{(N)}(\,\mathrm{d}t)$ samples a root t from the mass of the eigenvector, and the theorem states that the pair formed by the mass measure and the root converges in law to (μ_X, U) under $\overline{\mathbb{Q}}_X$. By the Cameron–Martin theorem, the random measure $p_{X,x}(t)\,\mathrm{d}t$ under $\mathbb{P}$ has the law of μ_X under $e^{g_X(x)}\,\mathrm{d}\mathbb{P}$. Indeed, tilting by $e^{g_X(x)}$ adds the drift $\sqrt{2X}$ to β on $[0,x]$, which reverses the drift of g_X to the left of x. The mass of an eigenvector rooted at x decays exponentially away from x in each direction, at rate X per unit length, up to the fluctuations of two independent Brownian motions. At large X, part (c) describes this shape on the scale $1/X$. (In the coupling given by Brownian scaling, the densities converge in L^1 almost surely.) Undoing the scaling, a distance s in (1.17) corresponds to $(4-E^2)W^2 s$ sites. Heuristically, at large fixed X, the mass of a selected eigenvector at distance m from its root is of order $\exp(\sqrt{2}\beta_s - |s|)$ with $s = m/((4-E^2)W^2)$.

These shapes coincide with those of the critical random Schrödinger operator. Under the identification of parameters stated after Theorem 1.10, and with the function S of Section 1.1, the pair formed by the density of Rifkind and Virág [109, Theorem 1.1] and its root U has the law of (μ_X, U) under $\overline{\mathbb{Q}}_X$, which is the limit in Theorem 1.12(a), and by Brownian scaling the density on $\mathbb{R}$ proportional to $S(4s)$ has the law of p_∞; we show this in Remark 12.7 and at the start of Section 12.5. Recentered at its center of mass, the same shape is the limit of the eigenfunctions of the continuous Anderson Hamiltonian in its crossover regime [35, Theorem 3]. We also determine the mean of the random shape.

Corollary 1.13. *Let p_∞ be as in* (1.17)*, and set $\bar{p}_\infty(s) = \mathbb{E}p_\infty(s)$ for $s \in \mathbb{R}$. Then $\bar{p}_\infty$ is even, $\bar{p}_\infty(0) = 1/3$, $\int_{\mathbb{R}} \bar{p}_\infty(s)\,\mathrm{d}s = 1$, and*

$$\bar{p}_\infty(s) = \frac{\pi^{7/2}}{32\, s^{3/2}} e^{-s/4}\big(1 + O(s^{-1})\big), \qquad s \to \infty. \tag{1.18}$$

Corollary 1.13 is proved in Section 12. Since $\beta_s/s \to 0$ almost surely, a typical sample of p_∞ satisfies $\log p_\infty(s) = -|s|(1+o(1))$ as $|s| \to \infty$, while by (1.18) the mean decays at rate $1/4$. The mean profile is therefore governed by rare samples, in which the Brownian motion is unusually large far from the root. The tail (1.18), with its constant, confirms a prediction of Efetov and Larkin [38, (56)] (see Section 1.3). The value $\bar{p}_\infty(0) = 1/3$

is half the slope in X of the limiting normalized mean inverse participation ratio in Theorem 1.15 below. Indeed, we show in Section 12 that $X^{-1}\mathcal{C}_X(x,x) \to \bar{p}_\infty(0)$ as $X \to \infty$, for all $x \in (0,1)$.

Local intensities and participation

We now turn to the local density of states at a single site. For $i \in \{1,\dots,N\}$, $\varepsilon > 0$, and $v \in \mathbb{R}$, we set

$$\ell^\varepsilon_{i,N}(v) = N \sum_{\alpha=1}^{N} w_\alpha(i) P_\varepsilon(x_\alpha - v).$$

We let $G(z) = (H-z)^{-1}$ denote the resolvent of H, for $z \in \mathbb{C}$ with $\operatorname{Im} z \neq 0$. Using the spectral decomposition $G_{ii}(z) = \sum_\alpha w_\alpha(i)/(\lambda_\alpha - z)$ and (1.3), we find

$$\ell^\varepsilon_{i,N}(v) = s_E^{-1} \operatorname{Im} G_{ii}\big(E + (v + \mathrm{i}\varepsilon)/a_N\big). \tag{1.19}$$

By Theorem 7.6(d), the mean of $G_{ii}(E + (v+\mathrm{i}\varepsilon)/a_N)$ converges to $-E/2 + \mathrm{i}s_E$, uniformly over sites i with i/N away from 0 and 1, and the expectation of $\ell^\varepsilon_{i,N}(v)$ therefore tends to one at such sites. The next theorem identifies the limit of the two-point function of the local density, at two energies and at two sites, for all pairs of interior sites.

Theorem 1.14. *Let $\delta \in (0,1/2)$. For $\varepsilon_1, \varepsilon_2 > 0$ and $u, v \in \mathbb{R}$, set $z = \pi(\varepsilon_1 + \varepsilon_2) - \mathrm{i}\pi(u-v)$. Uniformly in the critical regime, in $(\varepsilon_1, \varepsilon_2, u, v)$ in compact subsets of $(0,\infty)^2 \times \mathbb{R}^2$, and in sites i, j with $\delta \le i/N, j/N \le 1-\delta$, we have*

$$\mathbb{E}\ell^{\varepsilon_1}_{i,N}(u)\ell^{\varepsilon_2}_{j,N}(v) - 1 - \tfrac{1}{2}\operatorname{Re} K_{X_N(E)}(i/N, j/N; z) \longrightarrow 0 \qquad (i \neq j), \tag{1.20}$$

$$\mathbb{E}\ell^{\varepsilon_1}_{i,N}(u)\ell^{\varepsilon_2}_{i,N}(v) - 1 - \operatorname{Re} \widetilde{K}_{X_N(E)}(i/N; z) \longrightarrow 0, \tag{1.21}$$

and

$$N^2\mathbb{E} \sum_{\alpha=1}^{N} w_\alpha(i) w_\alpha(j) P_{\varepsilon_1}(x_\alpha - u) P_{\varepsilon_2}(x_\alpha - v) \\ - (1 + \mathbf{1}_{i=j})\, \mathcal{C}_{X_N(E)}(i/N, j/N)\, P_{\varepsilon_1+\varepsilon_2}(u-v) \longrightarrow 0. \tag{1.22}$$

Theorem 1.14 is proved in Section 11. The sum in (1.22) is the part of $\mathbb{E}\ell^{\varepsilon_1}_{i,N}(u)\ell^{\varepsilon_2}_{j,N}(v)$ that comes from single eigenvalues, and subtracting it from (1.20) or (1.21) gives the contribution of pairs of distinct eigenvalues. The theorem includes adjacent sites $j = i+1$, for which i/N and j/N tend to the same point x. The limit $1 + \frac{1}{2}\operatorname{Re} K_X(x,x;z)$ of the two-point function at adjacent sites differs from the limit $1 + \operatorname{Re}\widetilde{K}_X(x;z)$ at a single site. Skvortsov and Ostrovsky found the same distinction for an infinite wire, between nearby and coincident points, with the same weight $B/(B-F)$ at coincident points [126, (2), (7)–(8)].

The two kinds of limits are related through the behavior of the kernels at $z = 0$. In Theorem 10.4 and Section 11, we show that for $0 < x \le y < 1$,

$$\frac{z}{2} K_X(x, y; z) \longrightarrow \mathcal{C}_X(x, y), \qquad z\widetilde{K}_X(x; z) \longrightarrow 2\mathcal{C}_X(x, x), \qquad z \downarrow 0, \tag{1.23}$$

where z tends to zero along the positive real axis. In this sense, the eigenvector correlations are the residues of the kernels at $z = 0$. Since $\operatorname{Re} z^{-1} = P_{\varepsilon_1+\varepsilon_2}(u - v)$, the term $(1 + \mathbf{1}_{i=j})\, \mathcal{C}_X(x, y) \operatorname{Re} z^{-1}$ that comes from these residues is exactly the limit in (1.22). Our proofs use this relation in the reverse direction. In Sections 9 and 11, we extract the eigenvector correlations from the local densities at positive damping. The factor $1 + \mathbf{1}_{i=j}$ is the ratio $\mathbb{E}|\gamma|^4/(\mathbb{E}|\gamma|^2)^2 = 2$ for a centered complex Gaussian variable γ with $\mathbb{E}\gamma^2 = 0$. This suggests that, on the scale of single sites, the coordinates of an eigenvector behave like complex Gaussian variables whose variances are given by the mass measure.

The eigenvector correlations at the scale of single sites are given by the following theorem. For $E \in (-2, 2)$, $r > 0$, and $i, j \in \{1, \dots, N\}$, we define the *rooted mean profile*

$$P_N(E, r; i, j) = N \langle w_\alpha(j) \rangle^i_{E,r},$$

so that $N^{-1} P_N(E, r; i, j)$ is the mean mass at j of an eigenvector rooted at i. We let $\operatorname{IPR}(u_\alpha) = \sum_i w_\alpha(i)^2$ denote the inverse participation ratio of u_α.

Theorem 1.15. *Let $\delta \in (0, 1/2)$, and let $\bar{r}_N \to 0$ be a deterministic sequence. Uniformly in the critical regime and in $0 < r \le \bar{r}_N$,*

$$\sup_{\delta \le i/N,\, j/N \le 1-\delta} \left| P_N(E, r; i, j) - (1 + \mathbf{1}_{i=j})\, \mathcal{C}_{X_N(E)}(i/N, j/N) \right| \longrightarrow 0, \tag{1.24}$$

and the same holds with $P_N(E, r; i, j)$ replaced by $N^2 \langle w_\alpha(i) w_\alpha(j) \rangle_{E,r}$. Moreover, uniformly in the critical regime and in $0 < r \le \bar{r}_N$,

$$N \langle \operatorname{IPR}(u_\alpha) \rangle_{E,r} - 2 - \frac{2X_N(E)}{3} \longrightarrow 0, \tag{1.25}$$

and $\int_0^1 \mathcal{C}_X(x, x)\, \mathrm{d}x = 1 + X/3$ for all $X > 0$.

Theorem 1.15 is proved in Section 11. The profile (1.24) is a statement at the scale of single sites, with coefficient one for distinct sites, even when they are adjacent, and coefficient two for coincident sites. It does not follow from Theorem 1.10, which concerns weak limits of the mass measure. The limit (1.25) requires, in addition, bounds on the moments of the local densities at all sites, including those near the ends of the chain, which are part (c) of Theorem 7.6. Its proof applies to every nonnegative integrable window with compact support and positive integral in place of ϕ. By Definition 1.9, we have

$$\int_0^1 \mathcal{C}_X(x, x)\, \mathrm{d}x = \mathbb{E}_{\mathrm{Q}_X} \int_0^1 \Big(\frac{e^{g_X(t)}}{Z_X} \Big)^2 \mathrm{d}t.$$

The limit (1.25) confirms the prediction of Fyodorov and Mirlin [54, 56] (see Section 1.3). The participation ratio of a selected eigenvector is the product of the factor two from the scale of single sites and the participation ratio of the macroscopic profile. As $X \downarrow 0$, the limit $2 + 2X/3$ tends to 2, which is the limit of N IPR for the eigenvectors of the GUE (for these, $\mathbb{E}\,\mathrm{IPR} = 2/(N+1)$). For large X it grows like $2X/3$. Heuristically, a selected eigenvector then occupies of the order of N/X, or $(4 - E^2)W^2$, sites.

Endpoints and the metallic regime

The parameter X interpolates between the metallic regime $X \to 0$, where one expects the statistics of the GUE, and the localized regime $X \to \infty$, where one expects Poisson statistics and localized eigenvectors. The next proposition confirms this for the observables above. Recall that these limits are taken after the critical limit, as explained after Definition 1.2. We set

$$I_{\mathrm{GUE}}(z) = \frac{1 - e^{-2z}}{2z^2}, \qquad \Sigma^2_{\mathrm{GUE}}(L) = L - 2\int_0^L (L-s)\Big(\frac{\sin \pi s}{\pi s}\Big)^2 \,\mathrm{d}s,$$

and we let δ_{Leb} denote the point mass at the Lebesgue measure on $[0,1]$, viewed as a law on $\mathcal{P}([0,1])$.

Proposition 1.16. *Let $L \geq 0$ and $0 < x < y < 1$. As $X \downarrow 0$ and as $X \to \infty$, the objects in the first column of the following table converge to the entries of the second and the third columns, respectively.*

Object	$X \downarrow 0$	$X \to \infty$
pair measure $\nu_X(\,\mathrm{d}s)$	$\big(1 - (\sin \pi s/(\pi s))^2\big)\,\mathrm{d}s$	$\mathrm{d}s$
number variance $\Sigma^2_X(L)$	$\Sigma^2_{\mathrm{GUE}}(L)$	L
law of μ_X under $\mathbb{Q}_X$	δ_{Leb}	law of δ_U, with U uniform on $[0,1]$
eigenvector correlation $\mathcal{C}_X(x,y)$	1	0
kernel $\frac{1}{2}K_X(x,y;z)$	$I_{\mathrm{GUE}}(z)$	0

The pair measures converge vaguely, the laws converge weakly, and the kernels converge locally uniformly in $\mathrm{Re}\, z > 0$.

The proofs are in Sections 12 and 14. The first two rows identify the pair measure and the number variance of Ξ_X with those of the sine process (the local limit of the GUE) as $X \downarrow 0$, and with those of a Poisson process as $X \to \infty$. Recall from Corollary 1.7 that ν_X is determined by its Poisson transforms. One can check that $1 + \mathrm{Re}\, I_{\mathrm{GUE}}(\pi\varepsilon - \mathrm{i}\pi v)$ is the Poisson smoothing of the two-point function of the sine process, in the sense of (1.12). By Theorem 1.14, the last two rows say that at separated sites the two-energy intensity

tends to $1 + \operatorname{Re} I_{\mathrm{GUE}}(z)$ as $X \downarrow 0$ and to 1 as $X \to \infty$, while the contribution of single eigenvalues tends to $P_{\varepsilon_1+\varepsilon_2}(u-v)$ and to zero. We expect that Ξ_X converges to the sine process as $X \downarrow 0$ and to a Poisson process as $X \to \infty$. However, Proposition 1.16 concerns only the pair measures.

We now describe the first corrections to the metallic endpoint. We let

$$G_{\mathrm{N}}(x,y) = \frac{x^2+y^2}{2} - \max(x,y) + \frac{1}{3}, \qquad x,y \in [0,1],$$

denote the Green kernel of the Neumann Laplacian on $[0,1]$ acting on functions of mean zero. It satisfies $-\partial_x^2 G_{\mathrm{N}}(x,y) = \delta(x-y) - 1$, with $\partial_x G_{\mathrm{N}}(x,y) = 0$ at $x \in \{0,1\}$ and $\int_0^1 G_{\mathrm{N}}(x,y)\,\mathrm{d}x = 0$.

Proposition 1.17. *As $X \downarrow 0$,*

$$\mathcal{C}_X(x,y) = 1 + 2X G_{\mathrm{N}}(x,y) + O(X^2) \tag{1.26}$$

uniformly in $x,y \in [0,1]$, and

$$I_X(z) = I_{\mathrm{GUE}}(z) + \frac{X^2}{45}\bigl(1 - e^{-2z}\bigr) + O(X^3) \tag{1.27}$$

uniformly for z in compact subsets of $\{\operatorname{Re} z > 0\}$.

Proposition 1.17 is proved in Section 13. The first correction to the eigenvector correlation $\mathcal{C}_X$ of Corollary 1.11 is linear in X, while the first correction to I_X is quadratic. Both coefficients are spectral quantities of the Neumann Laplacian on $[0,1]$, whose nonzero eigenvalues are $(\pi n)^2$, with normalized eigenfunctions $\sqrt{2}\cos(\pi n x)$ for $n \geq 1$. Indeed,

$$G_{\mathrm{N}}(x,y) = \sum_{n\geq 1} \frac{2\cos(\pi n x)\cos(\pi n y)}{(\pi n)^2}, \qquad \frac{1}{45} = 2\sum_{n\geq 1}\frac{1}{(\pi n)^4} = 2\operatorname{Tr} G_{\mathrm{N}}^2.$$

By (1.12), the corresponding correction to the smoothed pair function $(P_\varepsilon * \nu_X)(v)$ is

$$\frac{X^2}{45}\operatorname{Re}\bigl(1 - e^{-2\pi\varepsilon + 2\pi \mathrm{i} v}\bigr) \longrightarrow \frac{2X^2}{45}\sin^2(\pi v), \qquad \varepsilon \downarrow 0.$$

This has the form of the correction to the Wigner–Dyson two-point function of a metallic wire obtained by Kravtsov and Mirlin [80, (15)], and the spatial term $2XG_{\mathrm{N}}$ appears in [94, (3.55), (3.74)]; we compare both terms with the physics literature in Section 1.3.

Fluctuations across scales

We finally consider the fluctuations of Ξ_X in windows of length $L \to \infty$, at fixed X. Recall from (1.14) that $\Sigma_X^2(L)$ is the variance of the number of points of Ξ_X in a window of length L. We set $f_L(s) = f(s/L)$ and $\Xi_X(f) = \sum_{x \in \Xi_X} f(x)$ for functions f on $\mathbb{R}$ and $L > 0$. We use the Fourier transform $\hat{f}(k) = \int_{\mathbb{R}} e^{-2\pi i k s} f(s)\,\mathrm{d}s$, and for functions f and g such that the integral converges absolutely we set

$$\langle f, g \rangle_{\dot{H}^{1/4}} = \int_{\mathbb{R}} |k|^{1/2} \hat{f}(k) \overline{\hat{g}(k)}\,\mathrm{d}k.$$

Theorem 1.18. *Let $X \in (0, \infty)$, and let $f, g \colon \mathbb{R} \to \mathbb{C}$ be compactly supported functions of bounded variation, or Schwartz functions. Then, as $L \to \infty$,*

$$\begin{aligned} \frac{\operatorname{Cov}\big(\Xi_X(f_L), \Xi_X(g_L)\big)}{\sqrt{XL}} &\longrightarrow \frac{1}{2\sqrt{\pi}} \langle f, g \rangle_{\dot{H}^{1/4}} \\ &= \frac{1}{16\pi^{3/2}} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\big(f(s) - f(s')\big)\overline{\big(g(s) - g(s')\big)}}{|s - s'|^{3/2}}\,\mathrm{d}s\,\mathrm{d}s', \end{aligned} \tag{1.28}$$

and both integrals converge absolutely. In particular, $\Sigma_X^2(L) \sim \pi^{-3/2}\sqrt{XL}$ as $L \to \infty$.

Theorem 1.18 is proved in Section 14. For the sine process, the variance of $\Xi(f_L)$ stays bounded for smooth f, and the number variance grows logarithmically; for a Poisson process of intensity one, these variances grow linearly. However, at every finite X, the process Ξ_X lies strictly between these cases. Its variances grow like $\sqrt{L}$, with the homogeneous Sobolev seminorm of order $1/4$ in place of the seminorm of order $1/2$ that describes the sine process. The asymptotics $\pi^{-3/2}\sqrt{XL}$ confirms the law of Altshuler and Shklovskii [5, (6)–(6a)], which they derived for a metal, $X \ll 1$. Physics predicts it at every X for windows much longer than X [90, (2), (25)–(26)], and for band matrices a version on mesoscopic scales was proved by Erdős and Knowles, also in the localized regime [42, Section 1.6] (see Section 1.3).

By Proposition 1.16, we have $\Sigma_X^2(L) \to L$ as $X \to \infty$ for all $L \geq 0$, while Theorem 1.18 implies that $\Sigma_X^2(L) = o(L)$ as $L \to \infty$ for all $X > 0$. In other words,

$$\lim_{X \to \infty} \lim_{L \to \infty} \frac{\Sigma_X^2(L)}{L} = 0, \qquad \lim_{L \to \infty} \lim_{X \to \infty} \frac{\Sigma_X^2(L)}{L} = 1.$$

The next theorem resolves this by describing the number variance on windows whose length is of order X.

Theorem 1.19. *Let I_0 denote the modified Bessel function of order zero. Let $m_\infty(u) = e^{-2/|u|} I_0(2/|u|)$ for $u \in \mathbb{R} \setminus \{0\}$, let $m_\infty(0) = 0$, and set*

$$V(\ell) = \int_{\mathbb{R}} m_\infty(u) \Big(\frac{\sin \pi \ell u}{\pi u}\Big)^2 \mathrm{d}u, \qquad \ell \geq 0. \tag{1.29}$$

Then $\Sigma_X^2(X\ell)/X \to V(\ell)$ *as* $X \to \infty$, *locally uniformly in* $\ell \in [0, \infty)$. *Further,* $V(\ell) \sim \ell$ *as* $\ell \downarrow 0$ *and* $V(\ell) \sim \pi^{-3/2}\sqrt{\ell}$ *as* $\ell \to \infty$, *and for* $\ell > 0$ *we have* $0 < V(\ell) < \ell$, *and* $V(\ell)/\ell$ *is strictly decreasing in* ℓ.

Theorem 1.19 is proved in Section 14, and Figure 1.2 shows the curve. The function m_∞ is the limit on the scale X of the Bartlett spectrum m_X of Ξ_X (the structure factor of physics), the positive measure whose Fourier transform is the reduced covariance measure $\delta_0 + \nu_X - \mathrm{d}s$ [22, Section 8.2]; we show in Section 14 that its Laplace transform $\int_{\mathbb{R}} e^{-2a|u|} m_\infty(u)\,\mathrm{d}u$ is the limit of $XI_X(Xa)$ as $X \to \infty$. In physics, m_∞ is the spectral form factor of a long unitary wire, with u the time in units of the Heisenberg time of a localization volume. In cold-atom quench experiments, it is the contrast of the peak of coherent forward scattering [91, (13)], and we compare Theorem 1.19 with the physics literature in Section 1.3. Theorem 1.19 proves these predictions for the limiting processes Ξ_X. We have $m_\infty(u) \to 1$ as $|u| \to \infty$, which gives Poisson behavior in short windows, and $m_\infty(u) \sim |u|^{1/2}/(2\sqrt{\pi})$ as $u \to 0$, which gives the law of Theorem 1.18 in long windows. The crossover has the following heuristic explanation. By Theorem 1.15, an eigenvector occupies about N/X sites when X is large, and the eigenvalues of eigenvectors localized in a common region of this size have unfolded spacings of order X. A window of length much smaller than X contains at most one eigenvalue from each region, and these eigenvalues are nearly independent. A window of length much larger than X contains many eigenvalues from each region, and these repel each other.

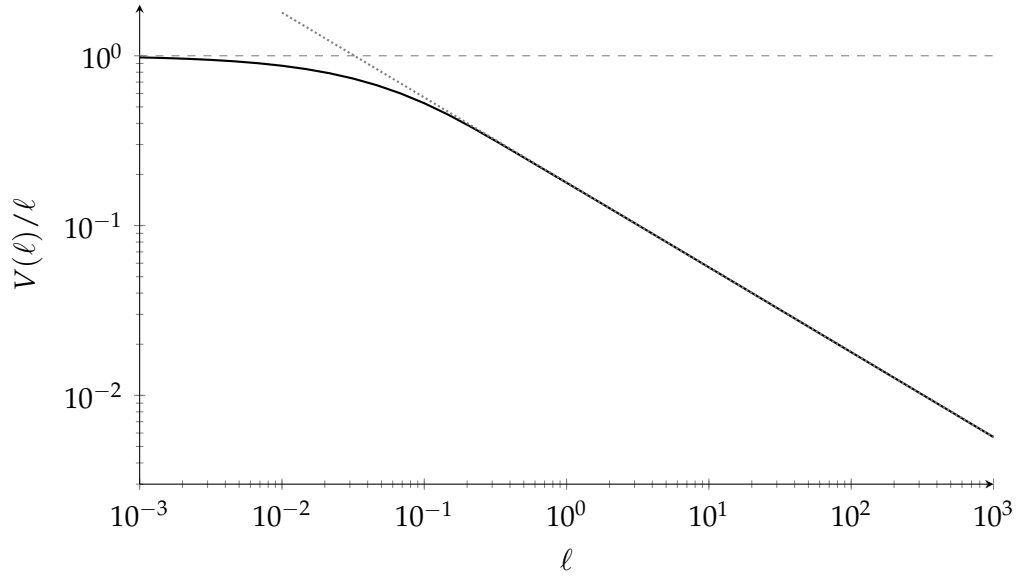


Figure 1.2. The variance per unit length $V(\ell)/\ell$ of Theorem 1.19 (solid line), computed by numerical quadrature with the supplementary script `crossover_ratio.py`, with the Poisson value 1 (dashed gray line) and the large-window asymptote $\pi^{-3/2}\ell^{-1/2}$ (dotted gray line).

Remark 1.20. We list the main questions that our results leave open. The results are proved only for $|E| \leq E_*$. The endpoint statements of Proposition 1.16 concern pair measures, number variances, and laws of the mass measure, and we do not prove that Ξ_X itself converges to the sine process or to a Poisson process. The crossover of Theorem 1.19 is a statement about the limiting processes Ξ_X, not about matrices with $X_N(E) \to \infty$. The two-energy intensities of Theorem 1.14 relate the eigenvalues to the eigenvectors, but we do not identify the joint limit of the eigenvalues and the mass measures as a marked point process. The eigenvector results describe one eigenvector selected by its eigenvalue; they do not give the law of a single coordinate $Nw_\alpha(i)$, or the joint law of several eigenvectors. Finally, Corollary 1.7 and Theorem 1.18 are second-moment statements, and we do not prove central limit theorems for linear statistics of Ξ_X. A point process is called number rigid if, for every bounded interval, the configuration outside the interval determines almost surely the number of points inside. Since the variance of $\Xi_X(f_L)$ grows like $\sqrt{L}$, the variance criterion for number rigidity of Ghosh and Peres [61, Theorem 6.1], which is stated for processes on $\mathbb{C}$ and whose proof applies on $\mathbb{R}$, cannot be met by dilations of a test function, and we do not know whether Ξ_X is number rigid.

1.3 Previous work and methods

Relation to the physics literature

The physics predictions for band matrices at the critical bandwidth come in two steps. The first is the reduction of the matrix model to the one-dimensional sigma model, which is a saddle-point approximation for wide bands without control of the errors. The second is the evaluation of observables within the sigma model, which is in many cases an exact computation and in others a perturbation expansion in X (see [94]). Our interface theorem, Theorem 7.6, makes the first step rigorous for the matrix of Definition 1.1 at $W^2 \asymp N$, and Part III carries out the second step rigorously for the observables of Section 1.2. (For the review [94], we use the equation numbers of the arXiv version.)

In the physics literature, a disordered wire in the unitary class is described by the ratio of the length of the wire to the length $\xi = 2\pi\nu A D_{\mathrm{diff}}$, which is given by the density of states ν, the cross-section A, and the diffusion constant D_{diff} [38, (52)–(53)] (see [94, Section 3.2.1]). For band matrices, Fyodorov and Mirlin wrote the sigma model on the lattice with a coupling equal to the square of the imaginary part of the saddle point times $\sum_r J(r)r^2$ [58, (7)]. Recall from (1.3) that $s_E = \sqrt{1 - E^2/4}$, so that $4s_E^2 = 4 - E^2$. For our profile, the saddle point, which solves the self-consistent equation of the semicircle law (see Section 2.3), has imaginary part s_E, and $\sum_r J(r)r^2 = 2W^2$ away from the ends of the chain, so that the coupling equals $2s_E^2 W^2$. Comparing it with the continuum

normalization of [94, (2.10)], we find that ξ is twice the coupling, and

$$\xi = (4 - E^2)W^2, \qquad X_N(E) = \frac{N}{\xi}.$$

The localization length given after [58, (7)] is 2ξ, four times the coupling; it is the decay length of $|\psi|$, while ξ is that of $|\psi|^2$. This identification is heuristic, but it agrees with the typical decay of the mass in Theorem 1.12, at rate one in units of $(4 - E^2)W^2$ sites, and with the parameter of the critical result [122, Theorem 1.2] on characteristic polynomials.

Several of our results are rigorous versions of predictions from the sigma model. The coarse law of Theorem 1.10 is the law of the envelope of an eigenfunction of a finite wire, which was derived by Fyodorov and Mirlin [56] and written by Mirlin as a path integral [94, (3.32)–(3.34)], following the method of Kolokolov for strictly one-dimensional systems [79]. Using the Girsanov theorem, one can check that the law [94, (3.34)] of the logarithm of the envelope is the law of $\log(e^{g_X}/Z_X)$ under $\mathbb{Q}_X$, with g_X, Z_X, and $\mathbb{Q}_X$ as in Definition 1.9; we write out this translation after (12.1). In the same description, an eigenfunction is the product of the envelope and a fast factor with Gaussian statistics [94, (3.33)], which accounts for the factor $1 + \mathbf{1}_{i=j}$ in Theorems 1.14 and 1.15. The law of Theorem 1.10 is also the eigenvector law of the critical one-dimensional random Schrödinger operator of Rifkind and Virág, by Remark 12.7. Band matrices are not among the operators of their universality conjectures [109, Conjectures 1.2–1.3], and Theorem 1.10 adds them to this class. The eigenvalue processes of the two models differ. As $X \downarrow 0$, the process Ξ_X tends to the sine process, in the sense of Proposition 1.16, while the eigenvalues of the weakly disordered Schrödinger operator are locally close to periodic [81].

The limit $2 + 2X/3$ of Theorem 1.15 agrees with the prediction of Fyodorov and Mirlin [54, 56]; its expansion to second order in X is in [57, (12), (17)]. Its slope $2X/3$ for large X agrees with [94, (3.48)]. The tail (1.18) of the mean profile, with the constant $\pi^{7/2}/32$, was obtained by Efetov and Larkin for the density correlator of a long wire [38, (56)], and the whole curve $\bar{p}_\infty$ agrees with the curve of [71, (75)–(76)], up to a factor two that comes from the normalization there (see Remark 12.13). The two-sided shape p_∞ of Theorem 1.12 is implicit in [94, (3.30)–(3.34)], and its counterparts for random Schrödinger operators are described after Theorem 1.12. The distinction between coincident and adjacent sites in Theorem 1.14, with the weight $B/(B - F)$ of the contact kernel, is the finite-wire counterpart of the local correlations of the infinite wire of Skvortsov and Ostrovsky [126, (2), (7)–(8)].

For the metallic corrections of Proposition 1.17, the spatial term $2XG_{\mathrm{N}}(x, y)$ corresponds to the diffusion propagator of a wire with reflecting ends [94, (2.22)–(2.23)], once the conductance in the normalization of [94] is identified with $1/X$ (see Remark 13.11). It appears in the first-order correlations of eigenfunction amplitudes [94, (3.55), (3.74)], derived in [9, 57]. The spectral term has the form of the correction of order g_{KM}^{-2} to

the Wigner–Dyson two-point function that Kravtsov and Mirlin obtained for a metallic sample of dimensionless conductance g_{KM} [80, (15)] (see also [94, (2.25), (2.27)]). The two corrections agree if $g_{\mathrm{KM}} = 1/(2\pi X)$. This identification is heuristic, and the agreement of the coefficients serves only to check consistency; we explain both in Remark 13.12.

The asymptotics $\Sigma^2_X(L) \sim \pi^{-3/2}\sqrt{XL}$ of Theorem 1.18, including the constant, is the formula of Altshuler and Shklovskii [5, (6)–(6a)] in dimension one and in the unitary class, since a window of L spacings corresponds to the ratio $2\pi XL$ of energy to Thouless energy in their notation. They derived it by perturbation theory for a good metal, which corresponds to $X \ll 1$, and for windows that are large compared with the Thouless energy. In long wires, physics predicts the same law for energy differences much larger than the level spacing of a localization volume [90, (2)], which in our units means windows of length much larger than X, and the formula [90, (25)–(26)] describes such windows at every value of the ratio of length to localization length. For band matrices, a rigorous version on mesoscopic energy scales, with the same seminorm $\int |k|^{1/2}\hat{f}(k)\overline{\hat{g}(k)}\,\mathrm{d}k$, was proved by Erdős and Knowles [42, Theorem 2.3] (see also [43]); as they note, it also holds in the localized regime [42, Section 1.6].

For the eigenvalue process, physics predicts a family of limits indexed by X, which tends to the GUE statistics as $X \to 0$, by Efetov's computation of the zero mode [37] (see [94, (2.11)]), and to Poisson statistics as $X \to \infty$. Theorem 1.6 identifies this family through its sources, and Proposition 1.16 confirms its two endpoints at the level of pair measures. For real symmetric block band matrices, the Poisson limit was proved directly in the localized regime in [47]. The sigma-model expression for the correlations of eigenfunction amplitudes at two energies separates the contribution of single eigenfunctions as a singular part [94, (3.4), (3.69)–(3.70)], which is the formal counterpart of the residues (1.23) and of the extraction in Section 9.

The structure factor m_∞ of Theorem 1.19 and the stationary pairing $\mathcal{I}$ of (14.26), which satisfies $\mathcal{I}(a) = \int_{\mathbb{R}} e^{-2a|u|} m_\infty(u)\,\mathrm{d}u$ by Lemma 14.12, were found in physics, for long unitary wires. The function m_∞ is the spectral form factor of such a wire. Micklitz, Müller, and Altland computed it as the contrast of the forward-scattering peak [91, (13)], and Marinho and Micklitz as the form factor itself [87, (19)]. The level correlation function of a long wire, whose transform is $\mathcal{I}$, was derived by Micklitz [90, (8)–(9)] and by Marinho and Micklitz [87, (12)–(13)]. In our variables, the latter formula is $X^{-1}\operatorname{Re}\mathcal{I}(-\mathrm{i}\pi s/X)$ at a separation of s mean level spacings, by (14.34), and Remark 14.13 compares the conventions and normalizations of these works with each other and with ours. We did not find the number variance V in these works. All of them use the exact zero mode of the singlet transfer operator of an infinite unitary wire, found by Skvortsov and Ostrovsky [126, (14)], which is also the stationary state in our proof in Section 14.

Methods

We give an overview of the proofs in Section 2. Here we describe the sources of their ingredients.

The supersymmetric method was introduced by Efetov [36, 37] (see [39, 133]). Accounts accessible to probabilists are [128], [7, Chapter 11], and [135], and Section 3 collects what we use. Supersymmetric transfer matrices were used early in rigorous work on one-dimensional random Schrödinger operators [14, 78] (see [29, Section 1]). For band matrices, the supersymmetric transfer-matrix formalism was suggested by Efetov and adapted at a heuristic level by Fyodorov and Mirlin [56] (see [119, Section 1]). Rigorous supersymmetric analyses of band matrices with the profile J of Definition 1.1 go back to the work of Disertori, Pinson, and Spencer on the density of states in three dimensions [27] (see also [25, 26]). In one dimension, a supersymmetric transfer operator was developed by M. Shcherbina and T. Shcherbina [115–117, 119, 122] (see the survey [118]) and by Disertori, Lohmann, and Sodin [29], who used a supersymmetry that commutes with the transfer operator and attribute it to Constantinescu [21]. Bao and Erdős [6] combined supersymmetry with comparison arguments for block band matrices. Our exact representation of $T_N^{(k)}$ as a transfer chain in Section 4 follows this line.

Our argument has three new steps, and Section 2.1 explains what each of them replaces and where it enters. First, we control the whole chain at $W^2 \asymp N$, including its noncompact sector, at every source rank k and with all coefficients in the Grassmann variables. At this bandwidth the chain converges to the heat semigroup of the sigma model over the whole interval $[0,1]$. We prove this by averaging over the fast variables of the chain, which relax over about W steps, much fewer than the N steps of the chain, with a normalization that an identity coming from supersymmetry (a Ward identity) makes exact. In the earlier work [118], the transfer operator is analyzed through its spectrum, above the threshold [119] and, at the threshold, for the compact sector alone [122]. Second, we attach both ends of the chain to its bulk up to errors $O(e^{-\gamma W})$, for some $\gamma > 0$. Third, we extract eigenvector moments from local densities, as described below.

For the noncompact sector, we use the hyperbolic Hubbard–Stratonovich transformation on the domain of Schäfer and Wegner [111], which was made rigorous by Zirnbauer [139, Section III.E, Lemma 3.2 and Theorem 3.3]. Other hyperbolic domains were treated by Fyodorov [52], by Fyodorov, Wei, and Zirnbauer [60], and by Müller-Hill and Zirnbauer [96], who also gave an explicit deformation of the Schäfer–Wegner domain into the Euclidean one [96, Appendix B]. These works concern a single matrix variable. The product contour for the chain, with all coefficients in the Grassmann variables, is constructed in Appendix B.

The normalization of the sigma model and the evaluation of the rank-one observables rest on supersymmetric localization (see Section 3.7), which goes back to Parisi and Sourlas [103]. Schwarz and Zaboronsky proved a general localization theorem on compact

supermanifolds [113, Theorem 1]. For the noncompact target space of our sigma model, we follow the argument of Disertori, Spencer, and Zirnbauer [28, Appendix C] (see also [7, Theorem 11.4.5]). On noncompact supermanifolds, changes of variables produce boundary terms, which were described by Rothstein [110] (see [139, Section II.A]) and are known in physics as Efetov–Wegner terms [76]. We work in global coordinates in which they do not arise; they reappear in the polar coordinates (B, F) as the mass that the singlet diffusion of Remark 1.5 carries into the corner.

The extraction of eigenvector moments from the part of Poisson-smoothed local densities that comes from coincident eigenvalues (Section 9) has, to our knowledge, no rigorous antecedent; its counterpart in physics is described in Section 1.3. Besides Theorem 7.6 and elementary properties of Gaussian matrices, Part III uses two known results about H, the factorial moment bounds of Peled, Schenker, Shamis, and Sodin [104, Theorem 3] and the density of states of [29, Theorem 1 and Remark 1.3]. The law of Ξ_X is determined through its Laplace functionals [23], which the ratios in (1.9) approximate by Fejér's theorem. The Brownian motion of Definition 1.9 arises from the noncompact sector. On the scale where B is large, the singlet diffusion moves in B approximately as a geometric Brownian motion. The physics counterpart of this scaling is the transfer equation [94, (3.22)–(3.23)], and our computation of the residues in Section 10 is a rigorous version of it. The resulting computations use Dufresne's identity [33], in the form of [88, Theorem 6.2], the Kontorovich–Lebedev transform [30, Section 10.43(v)], and the Liouville operator $-\partial_y^2 + e^y + 1/4$, whose use in one-dimensional localization goes back to Kolokolov [79] (see [94, (3.28)–(3.31)], and [20] for exponential functionals of Brownian motion in disordered systems). The function $2\sqrt{q}K_1(2\sqrt{q})$ that describes each side of a long wire is the zero mode of [38, (55)]. In Part IV, the passage from the chain to the continuum semigroup uses the Trotter–Kurtz approximation theorem [82, 131], in the form of [46, Chapter 1, Section 6], and the a priori bounds on partial chains use Hua's determinant inequality [68] (see [102]) and the Minami-type estimate of Graf and Vaghi [63].

Computations confirmed by computer

The calculations used in the proofs are written out explicitly in the text. Supplementary scripts check selected identities numerically or symbolically. Each script is named where it is used. We list four of the longer ones. The first two belong to the rank-one radial reduction of Section 6. They are the radial density d^{-2} of the invariant integral, which Lemma 6.2 establishes, and the radial part of C_1 in Lemma 6.8, which the proof of that lemma derives from the values of the Casimir operator in Lemma 5.11. Remark 6.9 names the two scripts that confirm them, the first through the divergence identity (6.20). This reduction is used in all of Part III, through K_X, $\widetilde{K}_X$, and I_X, but not in the proof of Theorem 7.6. The third is the form (10.22) of C_1 in the polar chart, which Lemma 10.7

proves from Lemmas 10.8 and 10.9, and which Remark 10.10 confirms. The fourth is the second-order metallic coefficient $I^{[2]}(z) = (1 - e^{-2z})/45$ in Lemma 13.8, the sum of the three integrals over the time simplex in (13.29), which Remark 13.9 confirms.

Organization of the paper

Section 2 gives an overview of the proof and ends, in Section 2.10, with a description of what each section of the four parts is for. Appendix A collects general tools, Appendix B constructs the product contour, and Appendix C proves the construction of the semigroup of the sigma model stated in Proposition 5.15.

1.4 Notation

We collect the conventions used throughout the paper. Symbols that are specific to one section are defined there.

Unless stated otherwise, $c, C > 0$ denote constants that may depend on the fixed parameters of the statement at hand, such as the source rank k, the damping ε, the margin δ, the bounds $X_\pm$ in Definition 1.2, and test functions, but not on N, W, or E. Their values may change from line to line. When the dependence on a parameter matters, we indicate it by a subscript, as in C_k. Let f be a complex-valued and g a nonnegative function of the parameters at hand. We write $f = O(g)$ if there exists a constant C such that $|f| \le Cg$, and $f = o(g)$ if $|f|/g \to 0$ in the limit under consideration; a subscript, as in $O_k(g)$, indicates a parameter on which the constant depends. For positive f and g, we write $f \asymp g$ if there exist constants $c, C > 0$ such that $cg \le f \le Cg$, and $f \sim g$ if $f/g \to 1$. In heuristic statements and in descriptions of other works, $f \ll g$ and $g \gg f$ mean that f is much smaller than g, in the sense of the work cited. Limits of quantities that depend on N are taken in the critical limit at X, or uniformly in the critical regime, as specified in Definition 1.2.

We set $[n] = \{1, \dots, n\}$ for positive integers n. As in Section 1.2, i denotes the imaginary unit, and we reserve the letters i and j for sites in $[N]$. We let Id denote the identity matrix or operator, $\mathbf{1}_A$ the indicator function of a set A, and δ_{ij} the Kronecker delta. For $a \in \mathbb{C}$, we let I_a and K_a denote the modified Bessel functions of order a; the functions I_X, K_X, and $\widetilde{K}_X$, indexed by the parameter X, are the kernels of Definition 1.4.

The matrix H is defined on a probability space with probability measure $\mathbb{P}$ and expectation $\mathbb{E}$. We use the same symbols for the law and the expectation of the Brownian motion β of Definition 1.9, which is defined on a separate probability space, and we let $\mathbb{E}_{\mathbb{Q}_X}$ denote the expectation under $\mathbb{Q}_X$. For complex random variables ζ_1 and ζ_2 with finite second moments, we set $\mathrm{Cov}(\zeta_1, \zeta_2) = \mathbb{E}[(\zeta_1 - \mathbb{E}\zeta_1)\overline{(\zeta_2 - \mathbb{E}\zeta_2)}]$ and $\mathrm{Var}(\zeta_1) = \mathrm{Cov}(\zeta_1, \zeta_1)$.

Recall that $G(z) = (H - z)^{-1}$ denotes the resolvent of H for $z \in \mathbb{C}$ with $\operatorname{Im} z \neq 0$, with entries $G_{ij}(z)$. By the spectral theorem,

$$G_{ij}(z) = \sum_{\alpha=1}^{N} \frac{u_\alpha(i)\overline{u_\alpha(j)}}{\lambda_\alpha - z},$$

and in particular $\operatorname{Im} G_{ii}(z) > 0$ for $\operatorname{Im} z > 0$. The supersymmetric variables of Parts II and IV have their own sign conventions, and Lemma 4.14 relates them to G. Grassmann algebras, supermatrices, the supertrace STr, and the superdeterminant SDet are introduced in Section 3.

Recall the Poisson kernel $P_\varepsilon(v) = \varepsilon/(\pi(v^2 + \varepsilon^2))$ and the Fourier transform $\hat{f}(k) = \int_{\mathbb{R}} e^{-2\pi i k s} f(s)\,\mathrm{d}s$ of Section 1.2. We have $P_{\varepsilon_1} * P_{\varepsilon_2} = P_{\varepsilon_1+\varepsilon_2}$. Point processes on $\mathbb{R}$ carry the vague topology, and spaces of probability measures carry the topology of weak convergence.

Supplementary scripts accompany the paper. They check identities numerically or in exact arithmetic, among them the computations listed in Section 1.3, and each script is named where it is used. The scripts are distributed with the source files of this paper, in the folder `anc`, and the paths of the scripts named in the text, such as `polar/check_polar_density.py`, are relative to that folder.

The sections of the four parts are numbered consecutively, and the appendices are lettered. Theorems, propositions, lemmas, corollaries, definitions, and remarks share one counter within each section, and equations are numbered within each section. A reference such as Theorem 7.6(b) refers to part (b) of a numbered statement.

SECTION 2

Overview of the proof

In this section, we explain the main steps of our proofs and the ideas behind them. The discussion is heuristic, and the precise statements and proofs are in Parts II–IV; Section 2.10 describes what each section of the paper does. We use the notation of Section 1 and the conventions of Section 1.4. This section can be read without any background in supersymmetry.

2.1 Strategy and main difficulty

All our results are derived from expectations of ratios of characteristic polynomials of H. Recall from (1.4) that the source ratio $T_N^{(k)}(\mathbf{u}, \mathbf{b})$ is the expectation of a ratio of $2k$ characteristic polynomials of H to $2k$ others, evaluated at distance of order $1/N$ from the energy E, above and below the real axis. We also use its derivatives in perturbations of the spectral parameters at finitely many sites, which are expectations of products of resolvent entries at these sites.

The proof has two halves, which are joined by the interface theorem, Theorem 7.6. In its part (a), it states that $T_N^{(k)}(\mathbf{u}, \mathbf{b})$ converges to the partition function $\mathcal{Z}_{X,\varepsilon}^{(k)}(\mathbf{u}, \mathbf{b})$ of the supersymmetric sigma model on $[0, 1]$ of Definition 1.3, as $N, W \to \infty$ with $X_N(E) \to X$, for sources near the base point $(\mathbf{u}, \mathbf{b}) = (0, \varepsilon \mathbf{1})$. Its part (b) states that the derivatives in perturbations at sites $i_1 < \cdots < i_m$ with $i_a/N \to x_a \in (0, 1)$ converge to the same sigma-model expression with an explicit function of the field Q inserted at each point x_a, also when neighboring sites are perturbed. We let $\kappa_E = E/2$. Then $-\kappa_E + \mathrm{i}s_E$ is the value at $E + \mathrm{i}0$ of the Stieltjes transform of the semicircle law. Parts (c) and (d) state that the moments of the diagonal resolvent entries $G_{ii}(z)$, at distance of order $1/N$ from E, are bounded uniformly over all sites, and that $\mathbb{E}G_{ii}(z) \to -\kappa_E + \mathrm{i}s_E$ at sites i with i/N away from 0 and 1. We prove the interface theorem in Part IV, and from it and from the sigma model we derive the eigenvalue and eigenvector statistics of Section 1.2 in Part III. Figure 2.1 shows how the sections depend on each other.

The theorem holds for every source rank k and every damping $\varepsilon > 0$, with two energies $0 < E_* < E_{\mathrm{out}} \leq 1/2$ that do not depend on k or ε. Its limits (a) and (b) hold up to the outer energy bound E_{out}, the energy up to which the estimates of Part IV hold. Its uniform statements (c) and (d) hold up to the inner energy bound E_*, and the margin between the two bounds absorbs the shift of the energy by $O(1/N)$ in their proof. Part III

also uses this margin to control the eigenvalues far from E, and our results are therefore stated for $|E| \leq E_*$. Remark 7.7 explains the two bounds.

The main difficulty is that three features occur together at $W^2 \asymp N$. First, eigenvalue statistics require inverse determinants at energies in both half planes, and these bring in a noncompact symmetry group, as we explain in Section 2.2. Characteristic polynomials, which are the subject of the earlier critical results recalled in Section 1.1, involve only the compact part of the sigma model. Second, the supersymmetric representation of $T_N^{(k)}$ is a one-dimensional chain of N matrix-valued variables, and at $W^2 \asymp N$ it does not reduce to an integral over finitely many variables. Its slow variables diffuse over a time of order one along the chain, and we must control the chain over all $N \asymp W^2$ steps, with errors that remain summable over that many steps. Third, eigenvector statistics require all derivatives in the local perturbations, including every coefficient in the anticommuting variables, and they appear only in the limit $\varepsilon \downarrow 0$ of the damping ε of (1.4), which comes after the limit $N \to \infty$.

Three steps of our argument are new, and we point out each where it is explained. The first is the control of the whole chain at $W^2 \asymp N$, including its noncompact part, at every source rank k and with all anticommuting coefficients, by averaging over its fast variables, in Section 2.4. It replaces the spectral analysis of the transfer operator in the earlier work recalled in Section 1.3, which concerns the regime above the threshold or, at the threshold, the compact part alone. The second is the attachment of the two ends of the chain with errors of order $e^{-\gamma W}$ for some $\gamma > 0$, in Section 2.5. It replaces arguments by weak convergence or compactness, which give no rate, and a rate is needed, since the comparison along the chain multiplies these errors by quantities that we control only up to $e^{o(W)}$. The third is the extraction theorem for eigenvector moments, in Section 2.7, which replaces the formal residue identities through which eigenvector statistics are computed in physics [94, (3.4) and (3.69)–(3.70)], and which has, to our knowledge, no rigorous antecedent.

2.2 The transfer chain

We begin by explaining why a ratio of determinants becomes a one-dimensional chain. A determinant is a Gaussian integral over anticommuting, or Grassmann, variables, and an inverse determinant is a Gaussian integral over commuting complex variables, provided that the Hermitian part of the quadratic form is positive definite. We write each of the $4k$ determinants of $T_N^{(k)}$ in this way, as in Section 3.1. In the resulting integral, H appears linearly in the exponent, and the normalizations of the commuting and the anticommuting Gaussians cancel. We may then take the average over H inside the integral, and since H is Gaussian, this average is the exponential of a quartic form in the integration variables, with the coefficients J_{ij}. We decouple this quartic form by a

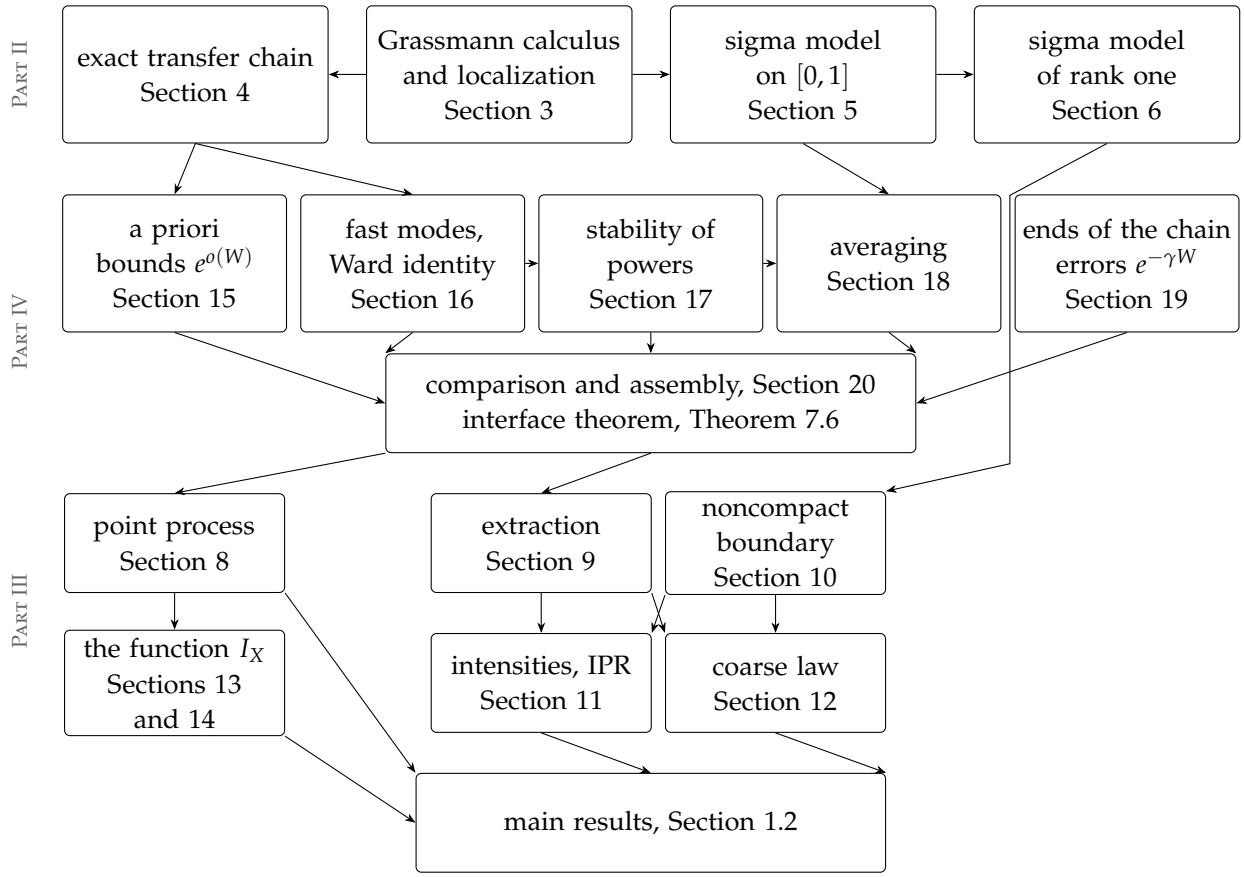


Figure 2.1. The logical structure of the proof. An arrow from one box to another means that the second uses results of the first, and only the main dependencies are drawn (for instance, Section 19 also uses results of Sections 16 and 18). Section 4 is used throughout Part IV, and Sections 5 and 6 throughout Part III, which uses from Part IV only the interface theorem. The a priori bounds $e^{o(W)}$ on partial chains (integrals over consecutive sites of the chain) make the errors $e^{-\gamma W}$, for some $\gamma > 0$, of the attachment of the ends negligible.

Hubbard–Stratonovich transformation (a Gaussian integral that turns the exponential of a quadratic form into an average of exponentials of linear forms), which introduces a supermatrix S_i of size $(2k|2k)$, in the sense of Section 1.2, for each site $i \in [N]$. The BB and FF blocks of S_i have commuting entries, while the BF and FB blocks have anticommuting entries.

The Gaussian field of this transformation has covariance J, so that its precision $J^{-1} = \mathrm{Id} - W^2\Delta$ appears in the integrand. This is where the variance profile matters. For all functions f on $[N]$, we have

$$\sum_{i,j=1}^{N} (J^{-1})_{ij} f(i) f(j) = \sum_{i=1}^{N} f(i)^2 + W^2 \sum_{i=1}^{N-1} \big(f(i+1) - f(i)\big)^2, \tag{2.1}$$

and the same identity holds for supermatrices, with products replaced by supertraces of products. The precision thus couples only neighboring sites, although J itself couples all pairs of sites. After the original variables are integrated out site by site, we obtain the exact representation of Proposition 4.11. To state it, we define the chain integral $\mathcal{Z}_N$ and the onsite weight w_Γ by

$$\begin{aligned} \mathcal{Z}_N(\Gamma) &= \int \prod_{i=1}^{N} w_\Gamma(S_i) \prod_{i=1}^{N-1} e^{-W^2 \operatorname{STr}(S_{i+1}-S_i)^2/2} \, \mathrm{d}S, \\ w_\Gamma(S) &= e^{-\operatorname{STr}(S^2)/2 - \mathrm{i} a_N^{-1} \operatorname{STr}(S\Gamma)} \operatorname{SDet}(E-S)^{-1}, \end{aligned} \tag{2.2}$$

where Γ is the source supermatrix of Section 1.2, and SDet is the superdeterminant of Section 3.4, which plays the role of the determinant for supermatrices. We let Γ_0 denote the value of Γ at the base point of the sources, at which $T_N^{(k)} = 1$. Up to an explicit factor that tends to one, $T_N^{(k)}$ is the ratio $\mathcal{Z}_N(\Gamma)/\mathcal{Z}_N(\Gamma_0)$, so that the normalization of the integral cancels, and by Theorem 7.6(a) this ratio converges to the partition function $\mathcal{Z}_{\mathrm{X},\varepsilon}^{(k)}$ of the sigma model. We have $\operatorname{STr}\Gamma = 0$, since $\sum_j u_j = 0$.

The integrand of (2.2) factors into terms that involve one site or two neighboring sites. The integral is a matrix element of T_Γ^{N-1}, where the transfer operator T_Γ of Definition 4.9 acts on functions of one supermatrix by applying the Gaussian link and multiplying the result by w_Γ. The link is the heat kernel on supermatrices at time $(2W^2)^{-1}$, and (2.2) resembles the Feynman–Kac representation of a walk whose Gaussian steps have size W^{-1}, in a potential. The analogy is only formal. The weights are complex and contain anticommuting variables, so that probabilistic arguments do not apply to the chain, and every estimate of Part IV is an estimate for operators.

The commuting variables need care, and this is where inverse determinants make a difference. We call the k pairs of determinants and inverse determinants with $\operatorname{Im} z > 0$ the retarded flavors, and the k pairs with $\operatorname{Im} z < 0$ the advanced flavors. The Gaussian integral that represents $\det(H-z)^{-1}$ has an exponent of the form $\mathrm{i}\zeta^*(z-H)\zeta$, with a complex vector ζ, when $\operatorname{Im} z > 0$, and the opposite sign is needed when $\operatorname{Im} z < 0$. The commuting variables of the $2k$ inverse determinants carry an indefinite quadratic form, with k positive and k negative directions, whose symmetry group is the noncompact group $U(k,k)$. In the Hubbard–Stratonovich transformation, this has the following consequence. Integrating the commuting block of each S_i over Hermitian matrices, as for the Gaussian unitary ensemble in Section 3.5, would produce an inverse superdeterminant with poles on the domain of integration. Following Schäfer and Wegner [111] (see Section 1.3), we integrate this block instead over a cycle in complex space, on which the commuting block of $E-S$, with its advanced rows multiplied by -1, has a positive definite imaginary part. We construct the product of these cycles for the whole chain, with all anticommuting coefficients, in Appendix B. The cycle is noncompact, and along its noncompact directions

the integrand decays only through the damping ε/a_N, which is of order W^{-2}. We will see in Section 2.7 that this slow decay is the origin of the Brownian motion in our eigenvector laws.

Local sources are treated in the same way. We call a supermatrix $Y = \mathrm{diag}(Y^{\mathrm{B}} \mid Y^{\mathrm{F}})$ that is added to $E + \mathrm{i}a_N^{-1}\Gamma$ at one site i a local source at i. It perturbs the spectral parameters at i and multiplies the onsite weight at i by $\mathrm{SDet}(\mathrm{Id} + (E - S_i)^{-1}Y)^{-1}$. The derivatives of this factor at $Y = 0$ are polynomials in the entries of $(E - S_i)^{-1}$, and the dictionary of Lemma 4.14 identifies their integrals with expectations of polynomials in the resolvent entries of H (for instance, a first derivative gives $-\mathbb{E}G_{ii}(z)$). We call the collection of all these derivatives at one site its complete local jet. The eigenvector observables need it in full, including its mixed and off-diagonal terms, since two derivatives at one site and two derivatives at neighboring sites have different limits.

2.3 The saddle manifold and the two scales

We now explain why the chain is close to a sigma model for large W. The source-free onsite weight w_0, obtained by setting $\Gamma = 0$ in w_Γ, is stationary at the supermatrices S with $S = (E - S)^{-1}$. For $S = \tau\mathrm{Id}$ with $\tau \in \mathbb{C}$, this equation reads $\tau(E - \tau) = 1$. Its solutions $\kappa_E \mp \mathrm{i}s_E$ are minus the semicircle values $-\kappa_E \pm \mathrm{i}s_E$ at $E \pm \mathrm{i}0$ recalled in Section 2.1, and the saddle-point equation is the self-consistent equation of the semicircle law. Since the equation is invariant under conjugation, its supermatrix solutions $\kappa_E\mathrm{Id} - \mathrm{i}s_E Q$, with $Q^2 = \mathrm{Id}$, form orbits. Recall the supermatrix $\Lambda = \mathrm{diag}(\Lambda_{\mathrm{B}} \mid \Lambda_{\mathrm{B}})$ of Section 1.2, whose diagonal entries are 1 on the retarded and -1 on the advanced flavors. The damping selects the orbit of the point $\kappa_E\mathrm{Id} - \mathrm{i}s_E\Lambda$, at which the retarded flavors take the value $\kappa_E - \mathrm{i}s_E$ and the advanced flavors the value $\kappa_E + \mathrm{i}s_E$. This orbit is a copy of the target space $\mathcal{M}_k$ of the sigma model, on which Q is the coordinate, and the onsite weight w_0 is identically one on it. In the FF block, the two values can also be distributed differently over the flavors. We call the union of all these orbits the saddle manifold, and we call the number $p \in \{0, \dots, 2k\}$ of diagonal entries of the FF block at which the value $\kappa_E - \mathrm{i}s_E$ occurs the signature of an orbit. The saddle manifold has $2k + 1$ components, one for each signature. The copy of $\mathcal{M}_k$ is the component $p = k$, the balanced component of Section 4.6, and we explain in Section 2.4 why the other components disappear in the limit.

The cycle of Section 2.2 passes through the saddle manifold, and by Lemma 15.9, the onsite weight on the cycle decays away from the saddle manifold. This is why the chain lives near the saddle manifold. Along the saddle manifold, the onsite weight is constant, while in the normal directions it has a nondegenerate critical point, with second derivatives of order one. The links $e^{-W^2\,\mathrm{STr}(S_{i+1}-S_i)^2/2}$ make the increments of the chain of size W^{-1}, and the two kinds of directions respond to them differently. We set $h = W^{-1/2}$,

so that the increment of one link is of size h^2. In the normal directions, the chain is a discretized Ornstein–Uhlenbeck process (a Gaussian autoregressive chain) with steps of size h^2 and a confinement of order one. Its normal coordinate relaxes by a factor $1 - ch^2$ per step, over about $h^{-2} = W$ steps. Its stationary size is of order h, since a chain with steps of size h^2 and relaxation rate h^2 has a stationary variance of order $h^4/h^2 = h^2$. In the tangent directions, the chain performs a random walk along the saddle manifold, each step of which has variance $\delta_W = h^4 = W^{-2}$ (the square of the increment of one link). Over the $N \asymp h^{-4}$ steps of the chain, it moves by a distance of order one, while over the h^{-2} steps that the normal coordinate needs to relax, it moves only by h. (In the language of physics, the normal and the tangent variables are the massive and the Goldstone modes.) Figure 16.1 shows the two scales.

The total time of the tangent walk is the parameter $X_N(E)$ of our results. Indeed, to first order, one link acts on functions of the tangent variables as the heat semigroup $e^{-t\mathsf{C}_k}$ at time $t = (4s_E^2W^2)^{-1}$, where C_k is the Casimir operator of Section 1.2, minus the invariant Laplace operator of $\mathcal{M}_k$. The reason is that the link factor, a Gaussian in the distance between neighbors, is the heat kernel of the ambient space at the time $(2W^2)^{-1}$, and Lemma 16.5 identifies the restriction of the ambient heat generator to the saddle manifold as $-(2s_E^2)^{-1}\mathsf{C}_k$. Since $4s_E^2 = 4 - E^2$, the N links of the chain run this semigroup for the time

$$\frac{N}{4s_E^2W^2} = \frac{N}{(4-E^2)W^2} = X_N(E). \tag{2.3}$$

The source factor $e^{-\mathrm{i}a_N^{-1}\operatorname{STr}(S\Gamma)}$ of an onsite weight becomes $e^{-s_Ea_N^{-1}\operatorname{STr}(Q\Gamma)}$ on the saddle manifold, since $\operatorname{STr}\Gamma = 0$. Since $Ns_E/a_N = \pi$, the N source factors combine to $\exp(-\pi\int_0^1 \operatorname{STr}(Q(t)\Gamma)\,\mathrm{d}t)$, where $Q(t)$ is the tangent variable at the site tN.

Heuristically, the chain should therefore converge to the sigma model of Section 1.2. We let $\mathsf{G}_{X,\Gamma} = X\mathsf{C}_k + \pi\operatorname{STr}(Q\Gamma)$ denote its generator. In this notation, $\mathcal{Z}^{(k)}_{X,\varepsilon}(\mathbf{u},\mathbf{b}) = \int e^{-\mathsf{G}_{X,\Gamma}}1$, with the invariant Berezin integral over $\mathcal{M}_k$ of Section 1.2. In probabilistic terms, this would be the expectation of $\exp(-\pi\int_0^1 \operatorname{STr}(Q(t)\Gamma)\,\mathrm{d}t)$ for a Brownian motion $Q(t)$ on $\mathcal{M}_k$ with generator $-X\mathsf{C}_k$, integrated over its starting point, where the time t is the position along the chain. The noncompact factor $U(k,k)/(U(k)\times U(k))$ of the underlying manifold of $\mathcal{M}_k$ comes from the inverse determinants, and the compact factor from the determinants. A Brownian motion on this superspace is only a heuristic notion, and Section 5 defines the semigroup $e^{-t\mathsf{G}_{X,\Gamma}}$ analytically, on weighted L^2 spaces. The weights are needed since the constant function 1, to which the semigroup is applied, is not square integrable over the noncompact factor.

The sigma model inherits the normalization of the source ratios, $\mathcal{Z}^{(k)}_{X,\varepsilon}(0,\varepsilon\mathbf{1}) = 1$. This is a consequence of supersymmetry, and since the same mechanism appears twice more below, we describe it. In flat space, there exists an odd vector field $\mathfrak{q}$, the supercharge of Section 3.7, which turns commuting variables into anticommuting ones and has zero

divergence. Consider a superfunction that $\mathfrak{q}$ annihilates. An integration by parts shows that inserting a factor $e^{-c\mathfrak{q}\sigma}$, where σ is an odd function for which $\mathfrak{q}\sigma$ is a Gaussian exponent in all variables, does not change its integral, for all $c > 0$. As $c \to \infty$, the inserted Gaussian concentrates the integral at the origin, where the Gaussian factors of the commuting and the anticommuting variables cancel. Therefore, the integral of the superfunction is its value at the origin, as in the dimensional reduction of Parisi and Sourlas [103]. We call a superfunction closed if a supercharge annihilates it. By the same argument, which is Lemma 5.19, the integral over $\mathcal{M}_k$ of a closed superfunction is its value at the base point $Q = \Lambda$. At the base point of the sources, the derivative in t of the integral of $e^{-t\mathsf{G}_{X,\Gamma_0}}1$ reduces, after an integration by parts, to the integral of a closed superfunction that vanishes at Λ. By localization, this derivative is zero, and we obtain the normalization in Corollary 5.20.

2.4 Averaging over the fast modes

We now describe how Part IV proves the convergence of the chain to the sigma model, which is the main part of the proof of Theorem 7.6(a). We write a supermatrix near the saddle manifold in terms of a point of the saddle manifold and a normal displacement U, in a tube of fixed small radius around the saddle manifold. We conjugate the transfer operator by the square root of the source-free onsite weight and by a Gaussian in U, so that the normal variables live in the Gaussian Hilbert space of a harmonic oscillator, with all their anticommuting coefficients. We let A_h denote the resulting operator, which we call the adapted transfer operator, as in Section 16.5.

We call the operator obtained from A_h by removing the tangent motion and the source the frozen normal transfer. By Lemma 16.10, it is to leading order the transfer operator of a supersymmetric harmonic oscillator, namely the Ornstein–Uhlenbeck semigroup at a complex time of order h^2, tensored with its finite-dimensional analogue for the anticommuting normal variables. We let Ω_0 denote the vacuum of the oscillator, which is the constant function in its Gaussian Hilbert space, and we call the number by which one step of the oscillator multiplies Ω_0 its zero-point factor. The commuting and the anticommuting normal coordinates contribute reciprocal factors to it, and in the balanced component, where the two kinds of coordinates have equal dimensions, these factors cancel. There the oscillator fixes Ω_0 and contracts the orthogonal complement of the vacuum by a factor $1 - ch^2$ per step.

Since the normal variables equilibrate after about h^{-2} steps, long before the tangent variables move, the evolution of the tangent variables should involve the normal ones only through their equilibrium. In the language of semigroup theory, this is averaging to first order over the fast variables, as in Kurtz's theorem on perturbed semigroups [83]. Rather than an averaging theorem, our argument uses the Trotter–Kurtz approximation

theorem [82, 131], and we check its two conditions, consistency on a suitable set of vectors and stability of the powers of A_h. The difficulty in both is the length of the chain. An error of size h^3 per step, which looks small, adds up to h^{-1} over the h^{-4} steps, and every error must be compared either with the gap h^2 of the oscillator or with the variance δ_W of the tangent step.

The first obstacle is the normalization of the fast chain. The frozen normal transfer differs from the oscillator by the anharmonic terms of the onsite weight, which connect the vacuum with excited states at order h^3 and change the vacuum itself at order h^4. A perturbative argument places its top eigenvalue at $1 + O(h^3 \cdot h^3/h^2 + h^4) = 1 + O(h^4)$. However, $(1 + Ch^4)^{h^{-4}}$ tends to e^C, so that such an argument would have to determine the coefficient of h^4 exactly, including all couplings of the anticommuting variables, to keep the normalization of the chain. To overcome this obstacle, we use supersymmetry, which shows that the relevant mean is preserved exactly. The normal oscillator has its own supercharge $\mathfrak{q}_{\mathrm{nor}}$, an odd operator with $\mathfrak{q}_{\mathrm{nor}}^2 = 0$. As in Section 2.3, we call the vectors that it annihilates closed. All factors of the frozen normal transfer are invariant, so that it commutes with $\mathfrak{q}_{\mathrm{nor}}$ and maps the closed vectors into closed vectors. For a closed vector, the normalized Gaussian integral over the normal variables is its value at the origin, by the localization argument of Section 2.3, and the invariant factors of the step equal one at the origin. It follows that the frozen normal transfer preserves the Gaussian mean of closed vectors exactly, whatever its anharmonic terms are. This is Proposition 16.22, and we call it the Ward identity, a name that the physics literature gives to all identities obtained from supersymmetry in this way.

Since the frozen normal transfer also contracts the closed vectors of mean zero by the factor e^{-ch^2} per step, Banach's fixed point theorem implies that there exists a unique closed vector Ω_h of mean one that it leaves invariant. By Lemma 18.3, it differs from the vacuum by $O(h)$, and every closed vector of mean one relaxes to it at the rate h^2 per step. Once the ends of the chain are attached to the tube, as in Section 2.5, they relax to Ω_h in this way. (A supersymmetry of the transfer operator was used by Disertori, Lohmann, and Sodin [29] to fix its top eigenvalue.)

We call the set of vectors $\Omega_h \otimes f$, where f is a function on the saddle manifold, the invariant normal line. The adapted step acts on it, to first order, as one step of the semigroup of the sigma model. Using the time normalization (2.3) in Proposition 18.12, we find that

$$A_h(\Omega_h \otimes f) = \Omega_h \otimes \big(f - N^{-1} \mathsf{G}_{X_N(E),\Gamma} f\big) + O(N^{-1}h) \tag{2.4}$$

for smooth functions f with compact support, with an error measured in norms of finitely many derivatives of f. The tangent heat of one link contributes the Casimir operator, and the source factor contributes the potential. This averaging on the invariant normal line, with its normalization made exact by the Ward identity, is the first new step of our argument. Over the N steps of a macroscopic time, the errors add up to $O(h)$, which tends

to zero. Heuristically, first order suffices because the tangent generator moves vectors off the invariant normal line only at order h, since its normal derivatives annihilate Ω_h, so that the correction of second order in the averaging vanishes in the limit.

The second obstacle is stability. The Trotter–Kurtz theorem also requires that the powers A_h^n with $n \leq Th^{-4}$ be bounded, uniformly in h, on all vectors and not only on the invariant normal line. It is natural to try to prove this by an energy estimate, showing that the generator of the adapted step is dissipative in the Gaussian Hilbert space of the normal oscillator. However, this approach fails for the anticommuting normal variables. Their part of the step has an eigenvalue of modulus at least $1 + ch^2$, which is compensated by the zero-point factor of the commuting variables alone, and Remark 16.11 explains why every metric of product form loses this compensation. To avoid this loss, we integrate the anticommuting normal variables against the complete Gaussian factor of the step before we take any norm.

We then compare A_h with a reference contraction C_h, the product of the oscillator and of the heat flow on the saddle manifold during one link. We bound $A_h - C_h$ relative to the defects $\mathrm{Id} - C_h^* C_h$ and $\mathrm{Id} - C_h C_h^*$, which measure how much C_h dissipates. A perturbation that is small relative to these defects keeps the powers bounded, by Lemma A.1. In this way, we obtain $\|A_h^n\| \leq e^{CT}$ for $n\delta_W \leq T$, which is Theorem 17.1 and the most technical estimate of Part IV. The comparison rests on the kernel of the step in coordinates centered at the midpoint of the link. In these coordinates, the Gaussian factor of a link is an even function of the tangent increment, and the terms of the kernel that are odd in the increment gain a factor h^2 when they are integrated.

With stability and consistency, the Trotter–Kurtz theorem implies that $A_h^n(\Omega_h \otimes f) \to \Omega_h \otimes e^{-t\mathsf{G}_{X,\Gamma}} f$ as $n/N \to t$, which is Theorem 18.14. The other components of the saddle manifold disappear in the same framework. In a component of signature $p \neq k$, the normal variables have $2(p-k)^2$ more commuting than anticommuting dimensions, with an excess of $(p-k)^2$ in each of the two root blocks. The zero-point factors of the oscillator then do not cancel, and they leave a factor of modulus $e^{-c(p-k)^2 h^2}$ per step. Over a stretch of θN steps, this gives the factor $e^{-c\theta W}$.

2.5 The ends of the chain

We now turn to the ends of the chain. The chain begins and ends with the constant function 1 on the cycle of Section 2.2, and this function is far from the invariant normal line in three ways. It does not decay in the normal variables. In the Gaussian space of the normal oscillator, it behaves like the inverse of the vacuum, and a decaying Gaussian factor is needed to give it a finite norm. It is spread over all $2k+1$ components of the saddle manifold. Finally, it lives on the cycle and not on the tube. The second new step of our argument is the attachment of the ends to the tube with exponentially small errors.

The errors must be exponentially small for the following reason. We call the chain on the cycle the physical chain, to distinguish it from the adapted chain on the tube, and we call $T_\Gamma^j 1$, the integral over j consecutive sites as a function of the variable at its free end, a partial chain. In the final comparison, every change from the cycle to the tube near an end is paired with a physical partial chain, and the size of these partial chains is not obvious. One physical step has norm at most $e^{C/W}$ in a suitable weighted space, which gives only e^{CW} over $N \asymp W^2$ steps. Observe that this is the same exponential order as the errors that Gaussian concentration produces, which are of size e^{-cWr^2} at normal distance r from the saddle manifold, since the normal fluctuations are of size $h = W^{-1/2}$. We need two estimates, an a priori bound $e^{o(W)}$ on the partial chains and attachment errors $e^{-\gamma W}$ for some $\gamma > 0$.

The a priori bound of Theorem 15.14, which is of order $e^{C\sqrt{W}(\log W)^2}$, combines two estimates of a different nature. Away from cores of radius $h \log W$ around the saddle manifold, one step contracts by the factor $e^{-c(\log W)^2/W}$, since the onsite weight decays there. Near the saddle manifold, we return to random matrices. The Gaussian part of a partial chain whose first site is linked to a given point S of the cycle can be integrated back, by reading the Hubbard–Stratonovich transformation backward. The resulting anchored identity, Proposition 4.13, writes the partial chain as the expectation, over an $n \times n$ Gaussian band matrix, of a quotient of two determinants, one from commuting and one from anticommuting variables, in which S appears linearly. Using Hua's determinant inequality [68], we compare the two determinants, and we bound the expectation by integrating over the diagonal entries of the band matrix, with the version of Minami's estimate [92] proved by Graf and Vaghi [63].

Next, we explain the attachment, which takes place in a block at one end of the chain, whose length is a universal constant times W. Since each link is a heat kernel at time $(2W^2)^{-1}$, the links of this block amount to a heat flow of total time of order $1/W$, which is short, so that a growth rate of order W per unit of heat time adds up to a constant factor over the block. This is why plain weighted energy estimates suffice at the ends, while the long stretches of the chain need the stability analysis of Section 2.4. Three such estimates are combined in Section 19. The first concerns the region far from the saddle manifold, where the constant function is not small but the onsite weight decays. An estimate of Agmon type [1] (an energy estimate with an exponential weight) shows that the partial chain over the block stays bounded against a weight that grows with the distance from the saddle manifold and is raised gradually during the block, and hence that it is exponentially small far from the saddle manifold.

The second estimate compares the block near the saddle manifold with reference vectors on the tube, which are built from adapted steps. The two satisfy the same local equation between steps and agree initially near the saddle manifold, so that their difference is created only at a normal distance of order one from it. A Davies–Gaffney

estimate [24] shows that a heat flow of duration of order $1/W$ carries a disturbance over a distance of order one only with a factor e^{-cW}, which yields the error $e^{-\gamma W}$. The third estimate gives the references a Gaussian decay in the normal variables, which the comparison along the rest of the chain needs, although at the start of the block they have finite norm only with a decaying Gaussian weight. Heuristically, since each step relaxes the normal coordinate by a factor $1 - ch^2$, the exponent of a Gaussian weight in the normal variables can be raised by an amount of order h^2 per step at a bounded cost, and over a stretch of order W steps it can move from negative to positive values.

One point requires care. In Theorem 7.6, the energy bounds E_* and E_{out} are fixed first, and the rank k, the damping ε, and the local sources of part (b) only afterward. The estimates that restrict the energy, at the ends of the chain and in the consistency analysis of Section 18, must therefore hold with constants that depend on none of these choices. These universal constants are not computed. They determine E_{out}, together with the restriction $|E| \leq 1/2$ of the a priori bound of Section 15, and this is why E_{out}, and with it the energy $E_* < E_{\mathrm{out}}$ of our results, is not explicit.

We compare the physical chain with the adapted one through a telescoping identity in one-sided form, Lemma 20.4. We attach the chain to the tube only at its right end, and we write $\mathcal{Z}_N(\Gamma)$ as the pairing of the constant function 1 with the chain applied to the attached block. In this form, each local error is paired with a physical partial chain started from 1, which the a priori bound controls, and the total error is of order $Ne^{-\gamma W + o(W)}$. Finally, Lemma 20.5 moves the pairing with the constant function from the cycle back to the tube by Stokes' theorem along an explicit homotopy, with an error of the same exponential order. On the tube, an initial layer of about $h^{-2}\log(1/h)$ adapted steps, whose tangent time tends to zero, relaxes the factors at both ends to $\Omega_h \otimes 1$, by the relaxation of closed vectors described in Section 2.4, and the unbalanced components are suppressed on stretches of length of order N. By the limit of the adapted chain on the invariant normal line, the balanced term converges to the integral over $\mathcal{M}_k$ of the product of $e^{-\theta \mathsf{G}_{X,\Gamma}}1$ and $e^{-(1-\theta)\mathsf{G}_{X,\Gamma}}1$, for every splitting point $\theta \in (0,1)$, and this integral is $\mathcal{Z}^{(k)}_{X,\varepsilon}(\mathbf{u},\mathbf{b})$. The normal factors contribute the normalized Berezin integral of Ω_h^2 over the normal variables, which Section 16.4 calls the analytic Berezin pairing of Ω_h with itself, and which is one by localization. This proves Theorem 7.6(a).

Local sources need one further idea. The physical comparison keeps exactly the complete local jet at each site that carries a local source. A complete jet changes the bounds only by a constant factor and a bounded number of polynomial weights, and it leaves the Gaussian decay unchanged, so that the estimates at the ends hold for blocks that contain such sites, with the same universal constants, also when such sites are adjacent. This is Lemma 20.11.

Only afterward, inside the adapted chain, whose powers are bounded, do we replace the jet by its value on the saddle manifold. We call $\mathcal{G}(Q) = -\kappa_E \mathrm{Id} + \mathrm{i}s_E Q$ the orbit

resolvent. On the tube, we have $(E-S)^{-1} = -\mathcal{G}(Q) + O(h)$. The retarded and advanced diagonal entries of $\mathcal{G}(\Lambda)$ are the semicircle limits $-\kappa_E \pm \mathrm{i}s_E$ of $G_{ii}(E \pm \mathrm{i}0)$. In this order of the two replacements, the algebraic error $O(h)$ is paired with the stable adapted chain and not with a physical partial chain, whose size is known only to be $e^{o(W)}$. We call the function obtained from the jet by replacing $(E-S)^{-1}$ with $-\mathcal{G}(Q)$ the orbit multiplier of the site, as in (7.17). In the limit, the jet acts as the multiplication by its orbit multiplier. This multiplier is a polynomial in the entries of $\mathcal{G}(Q)$. It acts trivially on the normal factor and therefore preserves Ω_h, and we obtain Theorem 7.6(b). Derivatives at distinct sites, even adjacent ones, become separate multipliers, while all derivatives at one site become a single multiplier, which contains the off-diagonal entries of $\mathcal{G}(Q)$.

Parts (c) and (d) of Theorem 7.6 follow from the same comparison. For the moments of G_{ii} at every site, also close to the ends, we attach the chain at the end farther from i, so that the stretch of the chain that arrives at i has length of order N, and we use the boundedness of the pairing with the adapted chain rather than its limit. The mean of G_{ii} at sites away from the ends is the case of one first derivative, which localization evaluates as $-\kappa_E + \mathrm{i}s_E$.

2.6 From the sigma model to the point process

We now turn to Part III, which uses Theorem 7.6 and the sigma model, and nothing else from Part IV. The convergence of Ξ_N in Theorem 1.6 rests on three properties. The bounds of [104] on factorial moments of eigenvalue counts, which are estimates of Minami type, give tightness, and together with the asymptotics of the expected density of eigenvalues from [29] they show that every limit is simple and has intensity one. Next, H has a Gaussian component proportional to the identity, independent of the rest of H, whose standard deviation after unfolding is at least of order $W^{1/2}$ by Lemma 7.4. A translation of the unfolded spectrum by a bounded amount therefore changes its law by $O(W^{-1/2})$, and every limit is stationary. Finally, the sources at one damping determine the law, by Fejér's theorem as explained after Theorem 1.6, since they determine the Laplace functionals of the process. Since Theorem 7.6(a) identifies $T_N^{(k)}$ only near the base point, we also use that $T_N^{(k)}$ is entire in the parameters $\mathbf{u}$ and $\mathbf{b}$, which appear only in the numerator of (1.4), with bounds that are uniform in N, and a normal-families argument (Montel's theorem) extends the convergence to all parameters.

The rank-one derivatives in Theorem 1.6(c) are where the singlet state and the kernels of Definition 1.4 come from. Superfunctions on $\mathcal{M}_1$ that are invariant under the stabilizer of Λ are functions of the two coordinates $(B, F) \in D$ of Section 1.2. By Lemma 6.8, the operator C_1 acts on them as $-d\Delta_D d^{-1}$, where $d = B - F$ and the conjugation by d comes from the four anticommuting variables. The generator reduces to $\mathsf{G}_{X,z} = X\mathsf{C}_1 + zd$, with $z = 2\pi\varepsilon$, at the base point of the sources. Thus $e^{-t\mathsf{G}_{X,z}}1$ is the probability that the singlet

diffusion of Remark 1.5 survives killing at rate zd up to time t, which is the singlet state $\Psi_t^{X,z}$. In the coordinates (B, F), the invariant integral over $\mathcal{M}_1$ consists of an integral over D and a point mass at the corner $(1, 1)$, which corresponds to Λ. We call this point mass the corner term. It is the Efetov–Wegner term of polar coordinates described in Section 3.6, and localization computes it. The second derivative in Theorem 1.6(c) is a sum of pairings with two insertions along the wire. Each insertion splits into a value at the corner, which the semigroup preserves by localization, and a part that evolves by the diffusion $(\tilde{B}_t, \tilde{F}_t)$ of Definition 1.4, and these parts produce the kernel K_X and the function I_X.

2.7 Eigenvector moments

Eigenvector moments are extracted from local densities at positive damping. Recall that the local density $\ell_{i,N}^{\varepsilon}$ of Section 1.2 is the Poisson smoothing of the measure $N \sum_\alpha w_\alpha(i) \delta_{x_\alpha}$. Let Π_N denote the expectation of the measure $N^2 \sum_{\alpha,\beta} w_\alpha(i) w_\beta(j) \delta_{(x_\alpha, x_\beta)}$ on $\mathbb{R}^2$. The part of Π_N with $\alpha = \beta$ lies on the diagonal, with a density that is the eigenvector moment we want. We have

$$\mathbb{E}\ell_{i,N}^{\varepsilon}(0)\ell_{j,N}^{\varepsilon}(0) = \int_{\mathbb{R}^2} P_\varepsilon(t) P_\varepsilon(v)\, \Pi_N(\mathrm{d}t, \mathrm{d}v), \qquad \int_{\mathbb{R}} P_\varepsilon(t)^2\, \mathrm{d}t = \frac{1}{2\pi\varepsilon}. \tag{2.5}$$

A diagonal part with constant density c therefore contributes $c/(2\pi\varepsilon)$. For the stationary limiting measures considered below, whose integrals against products of Poisson kernels are finite, the part without mass on the diagonal contributes $o(\varepsilon^{-1})$ as $\varepsilon \downarrow 0$. The density c is the residue at $z = 0$ of the limit at the damping $\varepsilon = z/(2\pi)$, and Theorem 7.6(b) expresses this limit through the sigma model.

To make this rigorous, we must deal with the order of limits, which is the main difficulty, since Theorem 7.6 identifies the limit $N \to \infty$ only at fixed $\varepsilon > 0$. To handle this, we consider vague limits of the expected measures Π_N themselves, and Lemma 9.5 establishes three properties of all such limits. It is stationary in the energy variable, by the Gaussian component of H. Its part from pairs of distinct eigenvalues puts no mass on the diagonal. This follows from the bound on second factorial moments of eigenvalue counts, which shows that the number of pairs of distinct unfolded eigenvalues in an interval of length r has expectation $O(r^2)$. One might expect that estimates on k eigenvalues are needed for the products over k sites that the coarse law involves, but the pair bound suffices at every order. Finally, Poisson kernels may be integrated against the limit. The Gaussian component of H also makes the expected diagonal measures Gaussian mixtures (convolutions with a centered Gaussian density) whose width tends to infinity, and for such mixtures vague convergence implies that the densities converge locally uniformly. Densities give the annealed selections over windows of every width $r \le \bar{r}_N$, including

windows much smaller than the spacing of the eigenvalues. This extraction theorem, Theorem 9.6, is the third new step of our argument.

It remains to compute residues in the sigma model, and this is where the noncompact sector appears. Consider the kernel $K_X(x, y; z)$ of Definition 1.4, whose residue at $z = 0$ is twice the eigenvector correlation $\mathcal{C}_X(x, y)$. It is built from diffusions on D killed at rate $z(B - F)$. For small z, the killing is negligible unless B is of order $1/z$, so that the relevant part of D lies far out in the noncompact direction, near the boundary at infinity of the hyperbolic plane. On this scale, the noncompact coordinate B_t of the singlet diffusion becomes a geometric Brownian motion. For the starting point $B_0 = Xq/z$, we have, approximately and in law as processes,

$$\frac{z}{X} B_t \approx q\, e^{g_X(t)}, \qquad z \int_0^1 B_t \, \mathrm{d}t \approx X q Z_X, \tag{2.6}$$

where t is the position along the wire, and $g_X(t) = \sqrt{2X}\beta_t - Xt$ and $Z_X = \int_0^1 e^{g_X(t)} \, \mathrm{d}t$ are the Brownian motion with drift and its exponential functional of Definition 1.9. The compact coordinate F, the angles and the anticommuting variables decouple on this scale, since their coupling to B is of order $1/B$ by Lemma 10.11.

The singlet state at time one therefore tends to $\mathbb{E}e^{-XqZ_X}$, the Laplace transform of the total mass. On this scale, an insertion at the position x multiplies by a factor proportional to B_x, and hence to $e^{g_X(x)}$. In these coordinates, the invariant measure of the hyperbolic plane becomes $\mathrm{d}B = Xz^{-1} \, \mathrm{d}q$, and the factor z^{-1} is the pole. Integrating over the starting point q and using

$$A^{1-k} = \frac{1}{(k-2)!} \int_0^\infty q^{k-2} e^{-qA} \, \mathrm{d}q, \qquad A > 0,\ k \geq 2, \tag{2.7}$$

we obtain the moments $M_k^X = \mathbb{E}[e^{g_X(x_1)+\cdots+g_X(x_k)} / Z_X^{k-1}]$ of Definition 1.9. Heuristically, the mass profile $e^{g_X(t)}/Z_X$ of a selected eigenvector is the profile of the noncompact coordinate B of the sigma model along the wire, on the scale $1/z$ on which the damping acts, divided by its integral. For $k = 2$, we couple the scaled diffusion with the geometric Brownian motion driven by the same Brownian motion, and this proves $zK_X(x, y; z)/2 \to \mathcal{C}_X(x, y)$ in Theorem 10.4.

For moments of order $k \geq 3$, the extraction needs a nonnegative observable of the resolvent whose part from single eigenvalues is proportional to the product $\prod_a w_\alpha(i_a)$ over k sites. We use permanents of $\operatorname{Im} G$ at these sites. Inverse determinants are Gaussian integrals over commuting variables, so that derivatives at k sites of a local source in the commuting block produce permanents, by Wick's theorem for complex Gaussians, just as derivatives of a determinant produce determinants. We can therefore stay at rank one for all orders. The expected permanent is proportional to the Poisson smoothing of a positive measure whose diagonal part is proportional to $\sum_\alpha \prod_a w_\alpha(i_a)\delta_{x_\alpha}$. The limit of

the expected permanent at fixed damping is a pairing with k source insertions in the rank-one sigma model. The source is not invariant, so that this limit needs the complete rank-one model, with its angular and anticommuting variables. The generator acts on the components of a superfunction, on the scale $B \asymp 1/z$, as the scalar singlet generator plus a matrix potential of order $1/B$. The source becomes B times a function of the angle of the hyperbolic plane and of the anticommuting variables, and these variables stay fixed in the limit. We then compute the residue, an explicit multiple of M_k^X, by a coupling argument in Lemma 10.12, which is the most technical step of Section 10.

The remaining steps of the coarse law are soft. By Scheffé's lemma, the moments at sites that are separated from each other and from the ends determine the step densities of all tensor moments of the mass measure in L^1, and by the Stone–Weierstrass theorem, these moments determine the law of the random probability measure on $[0,1]$, and Theorem 1.10 follows. The rooted laws of Theorem 1.12 follow from the Cameron–Martin theorem, as explained after that theorem, and the infinite-wire shape follows by Brownian scaling as $X \to \infty$. The mean shape of Corollary 1.13 is computed from Dufresne's identity for the exponential functional of a Brownian motion with drift [33].

2.8 Local intensities and the participation ratio

The local intensities of Theorem 1.14 involve two sites, and here the complete local jet matters. First derivatives at two distinct sites become two separate multipliers, each an entry of the orbit resolvent $\mathcal{G}(Q)$, and their pairing in the rank-one sigma model produces the kernel K_X. Two derivatives at one site become one multiplier, the corresponding term of the complete jet. Besides the product of two diagonal entries of $\mathcal{G}(Q)$, this multiplier contains the product of the two off-diagonal entries of the commuting block, which couple the retarded and the advanced flavors, as we explain in Remark 7.9. Its pairing produces the contact kernel $\widetilde{K}_X$, whose integrand carries the weight B/d. By (1.23), we have $zK_X/2 \to \mathcal{C}_X$ and $z\widetilde{K}_X \to 2\mathcal{C}_X$ as $z \downarrow 0$, and the extraction turns these residues into the profile of Theorem 1.15, with the coefficient $1 + \mathbf{1}_{i=j}$. The full two-energy intensity also needs the part of the pair measure away from its atom at zero separation. We recover it from its values at equal energies, since it is an even function of the separation.

The participation ratio $N\langle \mathrm{IPR}(u_\alpha)\rangle_{E,r}$ is the average over all sites i of $N^2\langle w_\alpha(i)^2\rangle_{E,r}$, including the sites near the ends of the chain, where Theorem 7.6(b) is not available. There we need only a uniform bound on the diagonal densities, which follows from the moment bounds of Theorem 7.6(c) at all sites. Strips of width δN at the ends then contribute $O(\delta)$, and the interior contributes $2\int_0^1 \mathcal{C}_X(x,x)\,\mathrm{d}x$. We evaluate this integral exactly, as $1 + X/3$, in Proposition 11.12, by a Laplace transform in the length of the wire and two integrated identities for the resolvent of the killed geometric Brownian motion.

2.9 Second-order statistics and fluctuations across scales

All second-order statistics of Ξ_X in Section 1.2 are read off from the function I_X of Definition 1.4, through two transforms of positive measures. Recall from Corollary 1.7 that the pair measure ν_X is determined by its Poisson transforms (1.12), which are expressed through I_X by Theorem 1.6(c) and stationarity. Moreover, the reduced covariance measure $\delta_0 + \nu_X - \mathrm{d}s$ of Ξ_X is positive definite, so that its Fourier transform is a positive measure. Following [22, Section 8.2], we call this measure, in the normalization of our Fourier transform, the Bartlett spectrum m_X of Ξ_X. In Lemma 14.4, we prove the Laplace identity

$$I_X(z) = \int_{\mathbb{R}} e^{-2z|k|}\, m_X(\mathrm{d}k), \qquad \operatorname{Re} z > 0.$$

Asymptotics of I_X therefore become statements about ν_X and m_X, through continuity theorems for these transforms.

As $X \downarrow 0$, the diffusion moves only by $O(\sqrt{X})$ during the times along the wire, and the semigroups in (1.6) are close to multiplications. We expand them in powers of X by Duhamel's formula. The difficulty is to bound the remainders uniformly over the closed simplex $0 \le x \le y \le 1$ of positions, since on its faces some of the times vanish and the corresponding factors give no decay in B. Recall that $d = B - F$. We overcome this difficulty by controlling the fluctuations of $\log(1+d)$ along the paths, which shows that the total killing along the wire produces decay faster than every power of B. The coefficients of the expansion are then elementary integrals over the simplex of positions and over D, and evaluating them, we obtain the expansion of I_X in Proposition 1.17. Afterward, Remark 13.11 notes the identity $1/45 = 2 \operatorname{Tr} G_{\mathrm{N}}^2$ of Section 1.2, which expresses the coefficient of X^2 through the eigenvalues of the Neumann Laplacian on $[0,1]$. The expansion of $\mathcal{C}_X$ is simpler, since it is a Gaussian integral in the parameter $\sqrt{2X}$.

As $X \to \infty$ with $z/X = a$ fixed, the wire is much longer than the time in which the diffusion relaxes, and $XI_X(Xa)$ is close to a stationary pairing $\mathcal{I}(a)$, which involves only the stationary singlet state of an infinite wire. This state is the zero mode found by Skvortsov and Ostrovsky [126], and $\mathcal{I}(a)$ is an explicit combination of modified Bessel functions. Inverting the Laplace transform gives the function m_∞ and the crossover of Theorem 1.19. These objects were found in physics from the same zero mode, as recalled in Section 1.3, and Theorem 1.19 makes these predictions rigorous, as statements about the limiting processes Ξ_X. At fixed X, the same estimate for large z determines the behavior of m_X near zero, which is the law $\Sigma_X^2(L) \sim \pi^{-3/2}\sqrt{XL}$ of Theorem 1.18.

2.10 Organization of the paper

In Part II, we introduce the supersymmetric tools and the two exact objects of the proof. Section 3 collects the facts about Grassmann variables, Berezin integrals, supermatrices,

and localization in flat space that the paper uses, for readers who are new to them. Section 4 writes the source ratios as the open transfer chain (2.2) on the product of Schäfer–Wegner cycles, together with the anchored identity for partial chains and the dictionary with resolvent entries; it is the starting point of Part IV. Section 5 defines the sigma model on $[0,1]$, constructs its semigroup, proves the localization identity and the normalization, and is used in Parts III and IV. Section 6 reduces the model of rank one to diffusions on D and to a calculus of rank-one pairings, from which the singlet state and the kernels of Definition 1.4 arise; Part III uses it throughout.

In Part III, we derive our main results from the sigma model and the interface theorem. Section 7 collects the inputs about H that do not involve supersymmetry and states the interface theorem, on which Part III relies as its only input from Part IV. Section 8 identifies the limiting point process and proves Theorem 1.6. Section 9 proves the extraction theorem, which turns limits at positive damping into eigenvector moments over small energy windows, for use in Sections 11 and 12. Section 10 computes the residues of the sigma-model kernels at zero damping through geometric Brownian motion at the noncompact boundary, and these residues identify the eigenvector moments. Section 11 proves Theorems 1.14 and 1.15, and Section 12 proves Theorems 1.10 and 1.12 and their corollaries. Section 13 proves the metallic expansion of Proposition 1.17, and Section 14 proves the statements on the pair measure, the large-scale covariance, and the crossover, with the remaining rows of Proposition 1.16.

The interface theorem is proved in Part IV. Section 15 proves the bounds on one physical step and the a priori bound $e^{o(W)}$ on partial chains, which makes errors of order $e^{-\gamma W}$ negligible in the comparison. Section 16 sets up coordinates near the saddle manifold and the fast normal oscillator, identifies in Lemma 16.5 the heat generator on the saddle manifold as a multiple of C_k, and proves the contraction of the oscillator and its Ward identity. Section 17 proves the stability of the powers of the adapted transfer operator over h^{-4} steps. Section 18 constructs the state Ω_h, deduces from Lemma 16.5 the time normalization (2.3) of one link, proves the consistency (2.4), and deduces the limit of the adapted chain for vectors of the form $\Omega_h \otimes f$. Section 19 attaches the ends of the chain to the tube with exponentially small errors. Section 20 combines these results into the proof of Theorem 7.6. Finally, Appendix A collects four general tools, Appendix B constructs the product contour of Section 4, and Appendix C contains the proof of the regular realization of the semigroup of Section 5.

PART II

Supersymmetry for random band matrices

Part II contains the calculus of Grassmann variables and Berezin integrals, the exact representation of the source ratios as the integral of a chain of supermatrices, and the sigma model on $[0, 1]$ in Section 5, with its reduction at rank one in Section 6. Section 2.10 describes what each section of Part II does, and Section 3 is written for readers who have not used anticommuting variables.

Part III uses the dictionary Lemma 4.14 between resolvent entries and insertions in the chain, and the saddle manifold of Section 4.6. A reader of Part III also needs the source supermatrices and local sources of Section 4.1, the Schäfer–Wegner cycle of Section 4.4, and the onsite weights and the chain integral of Definition 4.9, in terms of which the dictionary is stated. Part IV starts from the exact representation Proposition 4.11 and from the anchored identity Proposition 4.13 for partial chains. Both parts use the target space of Section 5, the partition function with its semigroups, and the localization identity. The sigma model of rank one of Section 6 is used throughout Part III and nowhere in Part IV.

A reader who takes the definitions on trust may skip the proofs in Section 3 and Section 5.2, and the proof of the regular realization in Appendix C. The start of Section 5 says which of its definitions a reader of Part III needs.

SECTION 3

Grassmann variables and Berezin integration

This section introduces the calculus of anticommuting variables that the rest of the paper uses. The supersymmetric method writes an average of a ratio of determinants of H as an integral over commuting and anticommuting variables, in which the average over H can be computed exactly. In Section 4, this gives the exact representation of the source ratio $T_N^{(k)}$ as an integral over supermatrices, and in Section 5 the same calculus describes the limiting sigma model. We write for a reader who has not used supersymmetry before.

The material is standard. Textbook accounts are given by Berezin [8], Efetov [39], Wegner [135], and, for mathematicians, Varadarajan [132]. Introductions aimed at probabilists are given in [128] and [7, Chapter 11], and Zirnbauer [140] reviews the method for Hermitian band matrices, with emphasis on the correctness of the mathematical reasoning. We prove the facts that later sections use, so that the paper can be read without these references. The one exception is the general change of variables formula, Theorem 3.15, for which we refer to the literature.

We begin in Section 3.1 by explaining why anticommuting variables appear. Section 3.2 defines Grassmann algebras and functions of even elements, and Section 3.3 introduces the Berezin integral and computes Gaussian integrals. Section 3.4 treats supermatrices, the supertrace, and the superdeterminant. With these tools, Section 3.5 carries out the whole method for a matrix from the Gaussian unitary ensemble, which shows what the calculus is for. Section 3.6 then discusses changes of variables and the boundary terms they can produce, and in Section 3.7, we prove the localization theorem in flat space; the same mechanism normalizes the sigma model in Section 5. The coefficient norms, with which Section 5 and Part IV estimate superfunctions, are defined in Section 5.2, where they are first used.

3.1 Motivation

Let A be an $N \times N$ complex matrix whose Hermitian part $\operatorname{Re} A = (A + A^*)/2$ is positive definite. We will show in Proposition 3.8 that

$$\int_{\mathbb{C}^N} e^{-\zeta^* A \zeta} \, \mathrm{d}\zeta = \frac{1}{\det A}, \qquad \int_{\mathbb{C}^N} \zeta_i \bar{\zeta}_j \, e^{-\zeta^* A \zeta} \, \mathrm{d}\zeta = \frac{(A^{-1})_{ij}}{\det A}, \tag{3.1}$$

where $\mathrm{d}\zeta$ is the Lebesgue measure on $\mathbb{C}^N$ divided by π^N. Let H be an $N \times N$ Hermitian matrix, let $z = E + \mathrm{i}\eta$ with $E \in \mathbb{R}$ and $\eta > 0$, and recall that $G(z) = (H - z)^{-1}$. We apply (3.1) with $A = \mathrm{i}(H - z)$, for which $\operatorname{Re} A = \eta \mathrm{Id}$. Since $A^{-1} = -\mathrm{i}G(z)$, we obtain

$$G_{ij}(z) = \mathrm{i} \frac{\int \zeta_i \bar{\zeta}_j \, e^{-\mathrm{i}\zeta^*(H-z)\zeta} \, \mathrm{d}\zeta}{\int e^{-\mathrm{i}\zeta^*(H-z)\zeta} \, \mathrm{d}\zeta}. \tag{3.2}$$

In this sense a resolvent entry is a covariance of a Gaussian weight on $\mathbb{C}^N$, although a complex one. For random H, the difficulty lies in the denominator of (3.2), which depends on H. Since the expectation of a ratio is not the ratio of the expectations, the average of $G_{ij}(z)$ cannot be computed by averaging the numerator and the denominator separately.

The denominator can be removed by a second Gaussian integral, over anticommuting variables $\chi_1, \bar{\chi}_1, \ldots, \chi_N, \bar{\chi}_N$, which produces the determinant instead of its inverse. Indeed, we will show that $\int e^{-\chi^* A \chi} \, \mathrm{d}\chi = \det A$ for every complex matrix A. Multiplying the two integrals, we will find in Lemma 3.9 that

$$\begin{aligned} \frac{\det(H - w)}{\det(H - z)} &= \int e^{-\mathrm{i}\zeta^*(H-z)\zeta - \mathrm{i}\chi^*(H-w)\chi} \, \mathrm{d}\zeta \, \mathrm{d}\chi, \\ G_{ij}(z) &= \mathrm{i} \int \zeta_i \bar{\zeta}_j \, e^{-\mathrm{i}\zeta^*(H-z)\zeta - \mathrm{i}\chi^*(H-z)\chi} \, \mathrm{d}\zeta \, \mathrm{d}\chi \end{aligned} \tag{3.3}$$

for all $w \in \mathbb{C}$. In these integrals, the normalizations of the two Gaussian integrals have canceled, and H enters linearly in the exponent. When H is Gaussian, its average can be taken inside the integral, where it gives an explicit exponential, and the average of a ratio of determinants becomes an integral of an explicit function. Derivatives in w, and products of such ratios at $E + \mathrm{i}\eta$ and $E - \mathrm{i}\eta$, generate products of resolvent entries.

This is the supersymmetric method. It was introduced by Efetov [36, 37] in the theory of disordered metals; superspace had been brought into statistical physics by Parisi and Sourlas [103], in the study of random magnetic fields. Verbaarschot, Weidenmüller, and Zirnbauer [133] applied the method to random matrix models of compound-nucleus scattering. The name refers to a symmetry between the commuting and the anticommuting variables of integrals such as (3.3), which we discuss in Section 3.7.

3.2 Grassmann algebras

We now define the algebras in which anticommuting variables take values. Throughout this section, $r \geq 0$ is an integer, and we write $[r] = \{1, \ldots, r\}$.

Definition 3.1. We define the *Grassmann algebra* $\bigwedge_r$ to be the complex associative algebra with unit generated by $\xi_1, \ldots, \xi_r$ subject to the relations $\xi_a \xi_b = -\xi_b \xi_a$ for all $a, b \in [r]$. For $I = \{i_1 < \cdots < i_s\} \subset [r]$, we set $\xi^I = \xi_{i_1} \cdots \xi_{i_s}$, with $\xi^\emptyset = 1$, and we call ξ^I the *exterior monomial* of degree $|I|$.

The monomials ξ^I form a basis of $\bigwedge_r$ (which is the exterior algebra of $\mathbb{C}^r$), and every $f \in \bigwedge_r$ has a unique expansion

$$f = \sum_{I \subset [r]} f_I \, \xi^I, \qquad f_I \in \mathbb{C}. \tag{3.4}$$

We call the numbers f_I the *coefficients* of f, the number $f_\emptyset$ the *body* of f, and $f - f_\emptyset$ the *soul* of f. We call f *even* if $f_I = 0$ whenever $|I|$ is odd, and *odd* if $f_I = 0$ whenever $|I|$ is even. We call f *homogeneous* if it is even or odd, and we assign the parity 0 to even and 1 to odd elements.

The relations imply $\xi_a^2 = 0$. One checks on monomials that $fg = (-1)^{\alpha\beta} gf$ for homogeneous f and g of parities α and β. In particular, even elements commute with all elements of $\bigwedge_r$, while odd elements anticommute with each other and have square zero. We call f *nilpotent* if $f_\emptyset = 0$. For nilpotent f, every monomial of f^s has degree at least s, and at least $2s$ if f is even. We conclude that

$$f^{r+1} = 0, \qquad \text{and} \qquad f^{\lfloor r/2 \rfloor + 1} = 0 \quad \text{if } f \text{ is even.} \tag{3.5}$$

It follows that f is invertible if and only if $f_\emptyset \neq 0$, and then $f^{-1} = f_\emptyset^{-1} \sum_{s=0}^{r} (-n/f_\emptyset)^s$ with $n = f - f_\emptyset$. The anticommuting variables of the physics literature are the generators ξ_a, and the commuting variables are complex numbers or, more generally, even elements.

We next allow the coefficients to depend on commuting variables. Let $m \geq 0$ and let $U \subset \mathbb{R}^m$ be open.

Definition 3.2. We define the space of *superfunctions* on the *superdomain* $U^{m|r}$ to be $C^\infty(U^{m|r}) = C^\infty(U) \otimes \bigwedge_r$. Its elements are the sums $f = \sum_I f_I \xi^I$ with complex-valued coefficients $f_I \in C^\infty(U)$, multiplied as in $\bigwedge_r$, with pointwise products of the coefficients. We call $x = (x_1, \dots, x_m) \in U$ the *even coordinates* and $\xi_1, \dots, \xi_r$ the *odd coordinates* of $U^{m|r}$.

The terms coefficient, body, soul, even, odd, and nilpotent are used for superfunctions as in Definition 3.1; the body $f_\emptyset$ of a superfunction is a function on U, and $f(x) \in \bigwedge_r$ denotes the value of f at $x \in U$. The spaces $C^k(U^{m|r})$ are defined in the same way, and so are holomorphic superfunctions on $U^{m|r}$ for open $U \subset \mathbb{C}^m$, whose coefficients are holomorphic.

Several formulas below contain odd elements that are not coordinates, such as the odd entries of a supermatrix or odd sources. We regard such elements as further generators of a larger Grassmann algebra $\bigwedge_{r+r'}$ containing $\bigwedge_r$, and we call these further generators *auxiliary*. Even coordinate derivatives and the integration over odd coordinates defined in Section 3.3 act coefficientwise in the auxiliary generators. Here we always write a superfunction as $f = \sum_I f_I \xi^I$ with the coefficients f_I, which may contain auxiliary generators, to the left of the monomials ξ^I. For instance, if θ is an auxiliary odd generator and $r = 1$, then the coefficient of ξ_1 in $\theta\xi_1$ is θ, and in $\xi_1\theta = -\theta\xi_1$ it is $-\theta$. The

odd coordinate derivatives defined below extend by graded linearity: for an auxiliary coefficient a of parity α, we have $\partial_{\xi_b}(ag) = (-1)^\alpha a\,\partial_{\xi_b} g$. In particular, $\partial_{\xi_1}(\theta\xi_1) = -\theta$.

For $i \in [m]$, the derivative ∂_{x_i} acts on the coefficients. For $a \in [r]$, every superfunction can be written uniquely as $f = f' + \xi_a f''$, where f' and f'' do not contain ξ_a, and we set $\partial_{\xi_a} f = f''$. One checks on monomials that $\partial_{\xi_a}\partial_{\xi_b} = -\partial_{\xi_b}\partial_{\xi_a}$ and that

$$\partial_{\xi_a}(fg) = (\partial_{\xi_a} f)g + (-1)^\alpha f\,\partial_{\xi_a} g$$

for homogeneous f of parity α. We call an operator $\mathcal{V} = \sum_i c_i \partial_{x_i} + \sum_a d_a \partial_{\xi_a}$, where c_i and d_a are superfunctions, a *vector field* on $U^{m|r}$. We call $\mathcal{V}$ even if the c_i are even and the d_a are odd, and odd if the c_i are odd and the d_a are even. A homogeneous vector field $\mathcal{V}$ of parity β satisfies $\mathcal{V}(fg) = (\mathcal{V}f)g + (-1)^{\alpha\beta} f\,\mathcal{V}g$ for homogeneous f of parity α.

Functions of even elements

Expressions such as $e^{-\zeta^* A\zeta}$, where A may have even entries, require functions of even elements. Since the soul of an even element is nilpotent, a Taylor expansion around the body terminates.

Definition 3.3. Let $a = (a_1, \dots, a_p)$ be even superfunctions on $U^{m|r}$, with bodies $a_\varnothing = (a_{1,\varnothing}, \dots, a_{p,\varnothing})$ and souls $n = a - a_\varnothing$. Let $V \subset \mathbb{R}^p$ be open, let $g \in C^\infty(V)$, and suppose that the bodies are real-valued with $a_\varnothing(U) \subset V$. We define

$$g(a) = \sum_{\alpha \in \mathbb{N}_0^p} \frac{(\partial^\alpha g)(a_\varnothing)}{\alpha!}\, n^\alpha, \qquad n^\alpha = n_1^{\alpha_1} \cdots n_p^{\alpha_p}. \tag{3.6}$$

If instead $V \subset \mathbb{C}^p$ and g is holomorphic on V, we define $g(a)$ by (3.6) with complex derivatives, for complex-valued bodies with $a_\varnothing(U) \subset V$.

By (3.5), the sum has only finitely many nonzero terms, those with $|\alpha| \le r/2$. If g is a polynomial, then Taylor's formula shows that $g(a)$ is the polynomial evaluated in the algebra. For complex coordinates $\zeta \in \mathbb{C}^m$, we use (3.6) with the real and imaginary parts as variables. In terms of the Wirtinger derivatives, a shift of ζ by even nilpotent elements u and of $\bar\zeta$ by even nilpotent elements v, which may be unrelated, reads

$$g(\zeta + u, \bar\zeta + v) = \sum_{\alpha,\beta} \frac{\partial_\zeta^\alpha \partial_{\bar\zeta}^\beta g(\zeta)}{\alpha!\,\beta!}\, u^\alpha v^\beta. \tag{3.7}$$

Lemma 3.4. *Let a, V, and g be as in Definition 3.3.*

(a) For all $\tilde g \in C^\infty(V)$, we have $(g\tilde g)(a) = g(a)\tilde g(a)$.

(b) For every homogeneous vector field $\mathcal{V}$ on $U^{m|r}$, we have $\mathcal{V}(g(a)) = \sum_{j=1}^p (\partial_j g)(a)\,\mathcal{V}a_j$.

(c) *Every identity between smooth, or holomorphic, functions on V remains true when the variables are replaced by even elements a with $a_\emptyset(U) \subset V$. In particular $e^{a_1+a_2} = e^{a_1}e^{a_2}$, $\det(e^K) = e^{\operatorname{Tr} K}$ for every square matrix K with even entries, and $g(a)^{-1} = (1/g)(a)$ if g has no zeros on V.*

(d) *Let $V' \subset \mathbb{R}^{p'}$ be open, let $\tilde{h} \in C^\infty(V; \mathbb{R}^{p'})$ with $\tilde{h}(V) \subset V'$, and let $g' \in C^\infty(V')$. Then $(g' \circ \tilde{h})(a) = g'(\tilde{h}(a))$, where $\tilde{h}(a)$ denotes the even elements $(\tilde{h}_1(a), \dots, \tilde{h}_{p'}(a))$.*

The same statements hold in the holomorphic case, and (d) also holds for complex-valued $\tilde{h}$ and g' holomorphic on an open set $V' \subset \mathbb{C}^{p'}$.

Proof. We begin with (a). We insert the two expansions (3.6) and collect the terms with a given power n^γ, which is possible because the souls are even and commute. By the Leibniz rule, we have

$$g(a)\tilde{g}(a) = \sum_\gamma \sum_{\alpha+\beta=\gamma} \frac{(\partial^\alpha g)(a_\emptyset)(\partial^\beta \tilde{g})(a_\emptyset)}{\alpha!\,\beta!} n^\gamma = \sum_\gamma \frac{\partial^\gamma (g\tilde{g})(a_\emptyset)}{\gamma!} n^\gamma = (g\tilde{g})(a).$$

We now prove (b). Both sides are left linear in $\mathcal{V}$ over superfunctions, and $(\partial_j g)(a)$ is even. It suffices to consider $\mathcal{V} = \partial_{x_i}$ and $\mathcal{V} = \partial_{\xi_b}$. Using the chain rule for the coefficients and the Leibniz rule for the commuting factors n_j, we obtain

$$\partial_{x_i} g(a) = \sum_\alpha \sum_{j=1}^p \frac{(\partial_j \partial^\alpha g)(a_\emptyset)}{\alpha!} (\partial_{x_i} a_{j,\emptyset})\, n^\alpha + \sum_\alpha \frac{(\partial^\alpha g)(a_\emptyset)}{\alpha!} \sum_{j=1}^p \alpha_j\, n^{\alpha - e_j}\, \partial_{x_i} n_j.$$

After the substitution $\alpha \mapsto \alpha + e_j$ in the second sum, each sum becomes an expansion of $(\partial_j g)(a)$, and we find

$$\partial_{x_i} g(a) = \sum_{j=1}^p (\partial_j g)(a)\, \partial_{x_i}(a_{j,\emptyset} + n_j) = \sum_{j=1}^p (\partial_j g)(a)\, \partial_{x_i} a_j.$$

The case $\mathcal{V} = \partial_{\xi_b}$ is the same computation without the first sum, since ∂_{ξ_b} annihilates the bodies and $\partial_{\xi_b}(n_j^{\alpha_j}) = \alpha_j n_j^{\alpha_j - 1} \partial_{\xi_b} n_j$ for even n_j.

Finally, if $g = \tilde{g}$ on V, then all derivatives agree, and (3.6) implies that $g(a) = \tilde{g}(a)$. This proves (c), and the examples follow with (a). For instance, $K \mapsto \det(e^K)$ and $K \mapsto e^{\operatorname{Tr} K}$ are entire functions of the entries of K that agree on complex matrices, and the determinant of a matrix with even entries is the same polynomial in the entries.

It remains to prove (d). We let $T_{b_0} u$ denote the Taylor polynomial of order $\lfloor r/2 \rfloor$ at b_0 of a smooth function u defined near b_0. The definition (3.6) says that $g(a) = (T_{a_\emptyset} g)(a)$, where the polynomial is evaluated in the algebra, since $n^\alpha = 0$ for $|\alpha| > r/2$. By the chain rule for Taylor polynomials, $T_{a_\emptyset}(g' \circ \tilde{h})$ is obtained from the composition $(T_{\tilde{h}(a_\emptyset)} g') \circ (T_{a_\emptyset} \tilde{h})$ by discarding the terms of degree larger than $\lfloor r/2 \rfloor$ in the variables centered at $a_\emptyset$. The discarded terms vanish at a, and therefore

$$(g' \circ \tilde{h})(a) = (T_{\tilde{h}(a_\emptyset)} g')\big((T_{a_\emptyset} \tilde{h})(a)\big) = (T_{\tilde{h}(a_\emptyset)} g')(\tilde{h}(a)) = g'(\tilde{h}(a)),$$

where the last step uses that $\tilde{h}(a)$ has the body $\tilde{h}(a_\emptyset)$ and a nilpotent soul. □

For real coefficients, the proof of [132, Theorem 4.3.1] defines $g(a)$, there called a pullback, by the expansion (3.6) and shows that $g \mapsto g(a)$ is a homomorphism of algebras, as in Lemma 3.4(a); see also [8].

3.3 Berezin integration

We now define integration over odd coordinates. The definition is algebraic, and the integral of a superfunction will be the ordinary integral of the coefficient of its top monomial.

Definition 3.5. For $f \in \bigwedge_r$, we define the *Berezin integral* $\int f \,\mathrm{d}\xi = f_{[r]}$ to be the coefficient of $\xi_1\xi_2\cdots\xi_r$ in f. We call a superfunction f on $U^{m|r}$ *integrable* if all its coefficients belong to $L^1(U)$, and for integrable f we define

$$\int_{U^{m|r}} f \,\mathrm{d}x \,\mathrm{d}\xi = \int_U f_{[r]}(x)\,\mathrm{d}x.$$

If f takes values in a Grassmann algebra with auxiliary generators, both integrals are taken coefficientwise in the auxiliary generators.

Exchanging two generators changes the sign of the Berezin integral, which is linear but not positive in any sense. Other references use other orders, which differ from ours by signs. Note that $\int f \,\mathrm{d}\xi$ is the body of $\partial_{\xi_r}\cdots\partial_{\xi_1} f$; in this sense the Berezin integral behaves more like a derivative than like an integral. We will need the following elementary properties; in (f), the odd variables $\theta_1, \dots, \theta_{r'}$ are integrated over as well.

Lemma 3.6. *Let $a \in [r]$ and $i \in [m]$.*

(a) For all $f \in \bigwedge_r$, we have $\int \partial_{\xi_a} f \,\mathrm{d}\xi = 0$. Consequently, for homogeneous f of parity α and all $g \in \bigwedge_r$,

$$\int (\partial_{\xi_a} f)\, g \,\mathrm{d}\xi = -(-1)^\alpha \int f\, \partial_{\xi_a} g \,\mathrm{d}\xi.$$

(b) Let f be a superfunction on $\mathbb{R}^{m|r}$ of class C^1 such that f and $\partial_{x_i} f$ are integrable. Then $\int \partial_{x_i} f \,\mathrm{d}x \,\mathrm{d}\xi = 0$.

(c) Let $B \in \mathbb{C}^{r\times r}$. For $f \in \bigwedge_r$, let $f(B\xi)$ denote the element obtained from (3.4) *by replacing each ξ_b with $\sum_c B_{bc}\xi_c$. Then $\int f(B\xi)\,\mathrm{d}\xi = \det B \int f \,\mathrm{d}\xi$.*

(d) Let $\theta_1, \dots, \theta_r$ be odd elements in the auxiliary generators, and let $f(\xi+\theta)$ denote the element obtained by replacing each ξ_b with $\xi_b + \theta_b$. Then $\int f(\xi+\theta)\,\mathrm{d}\xi = \int f \,\mathrm{d}\xi$.

(e) Let f be a superfunction on $\mathbb{R}^{m|r}$ whose coefficients are smooth and have all their partial derivatives integrable, and let $u_1, \dots, u_m$ be even nilpotent elements in the auxiliary generators. Then $\int f(x+u)\,\mathrm{d}x\,\mathrm{d}\xi = \int f \,\mathrm{d}x\,\mathrm{d}\xi$, where $f(x+u) = \sum_I f_I(x+u)\xi^I$ is defined by (3.6)*. The same holds on $\mathbb{C}^m$ for the independent shifts* (3.7) *of ζ and $\bar{\zeta}$.*

(f) Let $f \in \bigwedge_{r+r'}$, with generators $\xi_1, \dots, \xi_r, \theta_1, \dots, \theta_{r'}$. Then the iterated Berezin integrals in the two orders satisfy

$$\int \Big(\int f \, \mathrm{d}\xi \Big) \, \mathrm{d}\theta = (-1)^{rr'} \int \Big(\int f \, \mathrm{d}\theta \Big) \, \mathrm{d}\xi.$$

In this paper, r and r' are even whenever two Berezin integrals are exchanged, and the sign is one.

Proof. The top coefficient of $\partial_{\xi_a} f$ vanishes, since $\partial_{\xi_a} f$ does not contain ξ_a. Together with the Leibniz rule for $\partial_{\xi_a}(fg)$, this implies (a). Statement (b) is the identity

$$\int_{\mathbb{R}^m} \partial_{x_i} f_{[r]} \, \mathrm{d}x = 0$$

for an integrable C^1 function with integrable derivative, which follows from Fubini's theorem and the fundamental theorem of calculus.

We now prove (c). Only the term $f_{[r]} \prod_{b=1}^r (\sum_c B_{bc} \xi_c)$ of $f(B\xi)$ contains the top monomial, and

$$\prod_{b=1}^r \Big(\sum_{c=1}^r B_{bc} \xi_c \Big) = \sum_\pi \prod_{b=1}^r B_{b\pi(b)} \, \xi_{\pi(1)} \cdots \xi_{\pi(r)} = \det B \; \xi_1 \cdots \xi_r,$$

where the sum runs over the permutations π of $[r]$, and we used $\xi_{\pi(1)} \cdots \xi_{\pi(r)} = \operatorname{sgn}(\pi) \, \xi_1 \cdots \xi_r$. To prove (d), observe that $\prod_{b \in I} (\xi_b + \theta_b)$ contains the monomial $\xi_1 \cdots \xi_r$ only for $I = [r]$, and then with coefficient one.

It remains to prove (e). By the definition of $f_I(x+u)$ in Definition 3.3, we have

$$f_I(x+u) = \sum_\alpha \frac{\partial^\alpha f_I(x)}{\alpha!} \, u^\alpha,$$

and the terms with $\alpha \neq 0$ integrate to zero by (b). The shifts (3.7) are treated in the same way, since the Wirtinger derivatives are linear combinations of real derivatives.

Finally, with the coefficients written to the left, the left side of (f) is the coefficient of $\theta_1 \cdots \theta_{r'} \xi_1 \cdots \xi_r$ in f, and the integral $\int (\int f \, \mathrm{d}\theta) \, \mathrm{d}\xi$ is the coefficient of $\xi_1 \cdots \xi_r \theta_1 \cdots \theta_{r'}$. These two monomials differ by the sign $(-1)^{rr'}$. □

Statement (c) is the first sign that odd variables transform inversely to even ones: for even coordinates, the substitution $x \mapsto Bx$ multiplies the integral by $|\det B|^{-1}$. We will see in Section 3.6 that the two rules combine into the superdeterminant.

Gaussian integrals

The superspaces of this paper have complex even coordinates and pairs of odd coordinates.

Definition 3.7. Let $m, n \geq 0$. We write $\mathbb{C}^{m|2n}$ for the superdomain whose even coordinates $\zeta = (\zeta_1, \dots, \zeta_m) \in \mathbb{C}^m$ are identified with points of $\mathbb{R}^{2m}$, and whose odd coordinates are $\chi_1, \bar\chi_1, \dots, \chi_n, \bar\chi_n$, in this order. Here $\bar\zeta_j$ is the complex conjugate of ζ_j, while $\bar\chi_j$ is a generator independent of χ_j. We let $\mathrm{d}\zeta = \prod_{j=1}^m \pi^{-1} \,\mathrm{d}\operatorname{Re}\zeta_j \,\mathrm{d}\operatorname{Im}\zeta_j$, and for an integrable superfunction F on $\mathbb{C}^{m|2n}$ we set

$$\int F \,\mathrm{d}\zeta \,\mathrm{d}\chi = \int_{\mathbb{C}^m} F_{\mathrm{top}}(\zeta) \,\mathrm{d}\zeta,$$

where F_{top} is the coefficient of $\chi_1\bar\chi_1\chi_2\bar\chi_2 \cdots \chi_n\bar\chi_n$ in F. For $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$, we write $\zeta^* A\zeta = \sum_{i,j} \bar\zeta_i A_{ij} \zeta_j$ and $\chi^* B\chi = \sum_{i,j} \bar\chi_i B_{ij}\chi_j$, and we use the same notation for matrices with even entries. We let $\partial_{\zeta_j} = (\partial_{\operatorname{Re}\zeta_j} - \mathrm{i}\partial_{\operatorname{Im}\zeta_j})/2$ and $\partial_{\bar\zeta_j} = (\partial_{\operatorname{Re}\zeta_j} + \mathrm{i}\partial_{\operatorname{Im}\zeta_j})/2$ denote the Wirtinger derivatives.

The products $\chi_j\bar\chi_j$ are even, so the order of the pairs in the top monomial does not matter. (The factor π^{-1} in $\mathrm{d}\zeta$ is chosen so that $\int e^{-|\zeta|^2} \,\mathrm{d}\zeta = 1$ on $\mathbb{C}^m$.) Up to the normalization of $\mathrm{d}\zeta$, the integral $\int F \,\mathrm{d}\zeta \,\mathrm{d}\chi$ is the integral of Definition 3.5 on $\mathbb{R}^{2m|2n}$. We write $\operatorname{Re} A > 0$ for a complex matrix A if $\operatorname{Re} A = (A + A^*)/2$ is positive definite, and for a matrix A with even entries, we write $\operatorname{Re} A_\varnothing > 0$ if its body has this property. For a square matrix B, we let $\operatorname{adj} B$ denote its adjugate, so that $(\operatorname{adj} B)_{ij} = (-1)^{i+j} \det B_{[j,i]}$, where $B_{[j,i]}$ denotes B with row j and column i deleted. Recall that $\operatorname{adj} B = (\det B)B^{-1}$ if B is invertible.

Proposition 3.8. *Let $m, n \geq 1$, and let i and j be indices.*

(a) Let A be an $m \times m$ matrix with even entries such that $\operatorname{Re} A_\varnothing > 0$. Then $\det A$ is invertible, and

$$\int_{\mathbb{C}^m} e^{-\zeta^* A\zeta} \,\mathrm{d}\zeta = \frac{1}{\det A}, \qquad \int_{\mathbb{C}^m} \zeta_i\bar\zeta_j \, e^{-\zeta^* A\zeta} \,\mathrm{d}\zeta = \frac{(A^{-1})_{ij}}{\det A}.$$

(b) Let B be an $n \times n$ matrix with even entries. Then

$$\int e^{-\chi^* B\chi} \,\mathrm{d}\chi = \det B, \qquad \int \chi_i\bar\chi_j \, e^{-\chi^* B\chi} \,\mathrm{d}\chi = (\operatorname{adj} B)_{ij}.$$

(c) Let B be as in (b) and invertible, and let $\theta = (\theta_1, \dots, \theta_n)^T$ and $\bar\theta = (\bar\theta_1, \dots, \bar\theta_n)^T$ consist of odd elements in the auxiliary generators. Then

$$\int e^{-\chi^* B\chi + \chi^*\theta + \bar\theta^T\chi} \,\mathrm{d}\chi = \det B \; e^{\bar\theta^T B^{-1}\theta}.$$

(d) Let A be as in (a) and B as in (b). Then, on $\mathbb{C}^{m|2n}$,

$$\int e^{-\zeta^* A\zeta - \chi^* B\chi} \,\mathrm{d}\zeta \,\mathrm{d}\chi = \frac{\det B}{\det A}.$$

If $m = n$ and $B = A$, then $\int \zeta_i\bar\zeta_j \, e^{-\zeta^* A\zeta - \chi^* A\chi} \,\mathrm{d}\zeta \,\mathrm{d}\chi = \int \chi_i\bar\chi_j \, e^{-\zeta^* A\zeta - \chi^* A\chi} \,\mathrm{d}\zeta \,\mathrm{d}\chi = (A^{-1})_{ij}$.

Proof. We begin with (a) for complex A. The only step beyond linear algebra is the passage from Hermitian to non-Hermitian A, which we make by analytic continuation. We first let A be Hermitian and positive definite. We write $A = ODO^*$ with O unitary and $D = \operatorname{diag}(d_1, \dots, d_m)$. The substitution $\zeta = O\zeta'$ preserves $\mathrm{d}\zeta$, and

$$\int_{\mathbb{C}^m} e^{-\zeta^* A \zeta}\, \mathrm{d}\zeta = \prod_{j=1}^m \frac{1}{\pi} \int_{\mathbb{C}} e^{-d_j |\zeta'_j|^2}\, \mathrm{d}\operatorname{Re}\zeta'_j\, \mathrm{d}\operatorname{Im}\zeta'_j = \prod_{j=1}^m \frac{1}{d_j} = \frac{1}{\det A}.$$

Now let A be a complex matrix with $\operatorname{Re} A > 0$. We write $A = A_1 + \mathrm{i}A_2$, where $A_1 = \operatorname{Re} A$ and $A_2 = (A - A^*)/(2\mathrm{i})$ are Hermitian. The set $\mathcal{S}_A$ of $t \in \mathbb{C}$ with $\operatorname{Re}(A_1 + tA_2) > 0$ is a vertical strip containing 0 and i, because $\operatorname{Re}(A_1 + tA_2) = A_1 + (\operatorname{Re} t)A_2$. For $t \in \mathcal{S}_A$, the matrix $A_1 + tA_2$ is invertible, since $\operatorname{Re}(v^*(A_1 + tA_2)v) > 0$ for all $v \neq 0$. Moreover, for every compact set $\mathcal{K} \subset \mathcal{S}_A$ there exists $c > 0$ such that, for all $t \in \mathcal{K}$,

$$\left| e^{-\zeta^*(A_1 + tA_2)\zeta} \right| = e^{-\zeta^*(A_1 + (\operatorname{Re} t)A_2)\zeta} \leq e^{-c|\zeta|^2}.$$

Consequently the functions

$$t \mapsto \int_{\mathbb{C}^m} e^{-\zeta^*(A_1 + tA_2)\zeta}\, \mathrm{d}\zeta, \qquad t \mapsto \det(A_1 + tA_2)^{-1}$$

are holomorphic on $\mathcal{S}_A$. They agree for real $t \in \mathcal{S}_A$, where $A_1 + tA_2$ is Hermitian and positive definite, and the identity theorem implies that they agree at $t = \mathrm{i}$.

The second formula of (a) follows by differentiating the first in the entry A_{ji}, which is allowed by dominated convergence. Indeed, we have

$$\partial_{A_{ji}} e^{-\zeta^* A \zeta} = -\bar{\zeta}_j \zeta_i\, e^{-\zeta^* A \zeta}, \qquad \partial_{A_{ji}} \det(A)^{-1} = -\det(A)^{-1} (A^{-1})_{ij}.$$

Each side of the two formulas is a holomorphic function of A on the open set $\{A : \operatorname{Re} A > 0\}$, and their derivatives of all orders in the entries of A may be computed under the integral. For A with even entries, the integrand $e^{-\zeta^* A \zeta}$ is, by Definition 3.3, the Taylor expansion of $e^{-\zeta^* A' \zeta}$ in A' around $A_\varnothing$, evaluated at the soul $A - A_\varnothing$. Integrating term by term and using Lemma 3.4(c), we obtain (a) for even A, where $\det A$ is invertible because its body $\det A_\varnothing$ is nonzero.

We now prove (b). Each term of $\chi^* B \chi$ has degree two, and only the term $(-\chi^* B \chi)^n / n!$ of $e^{-\chi^* B \chi}$ contributes to the top monomial. Using $-\bar{\chi}_i \chi_j = \chi_j \bar{\chi}_i$, we have

$$\frac{(-\chi^* B \chi)^n}{n!} = \frac{1}{n!} \sum_{i_1, \dots, i_n} \sum_{j_1, \dots, j_n} \prod_{k=1}^n B_{i_k j_k}\, \chi_{j_k} \bar{\chi}_{i_k}.$$

A term is nonzero only if $(j_1, \dots, j_n)$ and $(i_1, \dots, i_n)$ are permutations of $[n]$. The factors $\chi_{j_k} \bar{\chi}_{i_k}$ are even, and ordering them by j_k accounts for the factor $n!$. A term then becomes

$\prod_l B_{\pi(l)l} \prod_l \chi_l \bar\chi_{\pi(l)}$ for a permutation π of $[n]$. Exchanging two of the $\bar\chi$ in the last product changes its sign, and we find

$$\prod_{l=1}^n \chi_l \bar\chi_{\pi(l)} = \operatorname{sgn}(\pi) \prod_{l=1}^n \chi_l \bar\chi_l, \qquad \left(e^{-\chi^* B \chi}\right)_{\text{top}} = \sum_\pi \operatorname{sgn}(\pi) \prod_{l=1}^n B_{\pi(l)l} = \det B.$$

The second formula follows by differentiating in B_{ji}, since

$$\partial_{B_{ji}} e^{-\chi^* B \chi} = -\bar\chi_j \chi_i \, e^{-\chi^* B \chi} = \chi_i \bar\chi_j \, e^{-\chi^* B \chi}, \qquad \partial_{B_{ji}} \det B = (\operatorname{adj} B)_{ij},$$

where the second identity is Jacobi's formula. Each side of these formulas is a polynomial in the entries of B, and the formulas remain valid for even entries.

To prove (c), we complete the square. Expanding the product and keeping the order of the odd factors, one checks that

$$-\chi^* B \chi + \chi^* \theta + \bar\theta^T \chi = -(\chi^* - \bar\theta^T B^{-1}) \, B \, (\chi - B^{-1}\theta) + \bar\theta^T B^{-1} \theta.$$

All terms are even, and by Lemma 3.4(c) the exponential factors accordingly. The first factor is obtained from $e^{-\chi^* B \chi}$ by shifting the generators χ_j and $\bar\chi_j$ by odd elements. By Lemma 3.6(d), its Berezin integral is that of $e^{-\chi^* B \chi}$, and the claim follows from (b).

Finally, using $e^{-\zeta^* A \zeta - \chi^* B \chi} = e^{-\zeta^* A \zeta} e^{-\chi^* B \chi}$ and (b), we find

$$\left(e^{-\zeta^* A \zeta - \chi^* B \chi}\right)_{\text{top}} = \det(B) \, e^{-\zeta^* A \zeta}.$$

Integrating over ζ with (a), we obtain (d). The two insertions follow in the same way from the second formulas of (a) and (b), since $(\operatorname{adj} A)_{ij} / \det A = (A^{-1})_{ij}$. □

We now apply Proposition 3.8 to resolvents, with the convention $G(z) = (H - z)^{-1}$. The sign in front of i in the exponent must be chosen so that the Gaussian integral over ζ converges, and this sign depends on the sign of $\operatorname{Im} z$.

Lemma 3.9. *Let H be an $N \times N$ Hermitian matrix, let $E \in \mathbb{R}$ and $\eta > 0$, and set $z_\pm = E \pm \mathrm{i}\eta$. Then, on $\mathbb{C}^{N|2N}$, for all $w \in \mathbb{C}$,*

$$\frac{\det(H - w)}{\det(H - z_\pm)} = \int e^{\mp \mathrm{i}[\zeta^*(H - z_\pm)\zeta + \chi^*(H - w)\chi]} \, \mathrm{d}\zeta \, \mathrm{d}\chi, \tag{3.8}$$

and for all $i, j \in [N]$,

$$\begin{aligned} G_{ij}(z_\pm) &= \pm \mathrm{i} \int \zeta_i \bar\zeta_j \, e^{\mp \mathrm{i}[\zeta^*(H - z_\pm)\zeta + \chi^*(H - z_\pm)\chi]} \, \mathrm{d}\zeta \, \mathrm{d}\chi \\ &= \pm \mathrm{i} \int \chi_i \bar\chi_j \, e^{\mp \mathrm{i}[\zeta^*(H - z_\pm)\zeta + \chi^*(H - z_\pm)\chi]} \, \mathrm{d}\zeta \, \mathrm{d}\chi. \end{aligned} \tag{3.9}$$

Proof. We set

$$A = \pm\mathrm{i}(H - z_\pm) = \pm\mathrm{i}(H - E) + \eta\mathrm{Id}.$$

Since $\pm\mathrm{i}(H - E)$ is anti-Hermitian, we have $\operatorname{Re} A = \eta\mathrm{Id} > 0$. Applying Proposition 3.8(d) with this A and with $B = \pm\mathrm{i}(H - w)$, we obtain

$$\int e^{-\zeta^* A\zeta - \chi^* B\chi}\,\mathrm{d}\zeta\,\mathrm{d}\chi = \frac{(\pm\mathrm{i})^N \det(H - w)}{(\pm\mathrm{i})^N \det(H - z_\pm)},$$

which is (3.8). For (3.9), we take $w = z_\pm$, so that $B = A$, and we use

$$(A^{-1})_{ij} = (\pm\mathrm{i})^{-1} G_{ij}(z_\pm) = \mp\mathrm{i}\, G_{ij}(z_\pm). \qquad \square$$

Multiplying a copy of (3.8) with z_+ by a copy with z_-, we write a product of two ratios of determinants as one integral over $\mathbb{C}^{2N|4N}$, with the signs $-\mathrm{i}$ and $+\mathrm{i}$ in the exponent. In Section 4, these signs are collected in the matrix Λ, and the dictionary between resolvent entries and Gaussian insertions is Lemma 4.14. Derivatives of (3.8) in w also generate resolvents; for instance, Jacobi's formula gives

$$\partial_w \frac{\det(H - w)}{\det(H - z)}\Big|_{w=z} = -\operatorname{Tr} G(z).$$

3.4 Supermatrices

Linear changes of all coordinates of a superspace, and the quadratic forms in Gaussian superintegrals, are described by supermatrices. Throughout this section, $\mathcal{R}$ denotes a Grassmann algebra, possibly with auxiliary generators, or an algebra of superfunctions. The even part of $\mathcal{R}$ is commutative, and determinants of matrices with even entries are defined by the usual formula.

Definition 3.10. Let $m, n \geq 0$. A *supermatrix* of size $(m|n)$ over $\mathcal{R}$ is a block matrix

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

with blocks $A \in \mathcal{R}^{m\times m}$, $B \in \mathcal{R}^{m\times n}$, $C \in \mathcal{R}^{n\times m}$, and $D \in \mathcal{R}^{n\times n}$. We call M even if the entries of A and D are even and those of B and C are odd, and odd if the entries of A and D are odd and those of B and C are even. We call A, B, C, and D the BB, BF, FB, and FF blocks of M. For homogeneous M of parity α, we define the *supertrace* $\operatorname{STr} M = \operatorname{Tr} A - (-1)^\alpha \operatorname{Tr} D$. For even M with A and D invertible, we define the *superdeterminant*, or *Berezinian*,

$$\operatorname{SDet} M = \frac{\det(A - BD^{-1}C)}{\det D}.$$

(The letters B and F stand for bosonic, or commuting, and fermionic, or anticommuting.) An even supermatrix maps columns $(v, \theta)^T$, whose first m entries v are even and whose last n entries θ are odd, to columns of the same kind. Products of supermatrices are matrix products, and the product of homogeneous supermatrices of parities α and β is homogeneous of parity $\alpha + \beta \pmod 2$. An even supermatrix M has the body $M_\emptyset = \operatorname{diag}(A_\emptyset, D_\emptyset)$. Suppose that $M_\emptyset$ is invertible, and suppose that $\mathcal{R}$ has r generators. Then $M = M_\emptyset(\mathrm{Id} + K)$, where $K = M_\emptyset^{-1}(M - M_\emptyset)$ has nilpotent entries and therefore satisfies $K^{r+1} = 0$, and

$$M^{-1} = \sum_{s=0}^{r} (-K)^s M_\emptyset^{-1}.$$

Conversely, the body of the identity $MM^{-1} = \mathrm{Id}$ shows that $M_\emptyset$ is invertible if M is. We conclude that an even supermatrix is invertible if and only if $A_\emptyset$ and $D_\emptyset$ are invertible; in this case $\operatorname{SDet} M$ is defined.

The following identities are standard, and their proofs are elementary. The only point at which the anticommutation of the odd entries matters is (3.10) below, which replaces the identity $\det(\mathrm{Id} - FF') = \det(\mathrm{Id} - F'F)$ for complex matrices F and F' by $\det(\mathrm{Id} - FF') = \det(\mathrm{Id} - F'F)^{-1}$ for matrices F and F' with odd entries.

Lemma 3.11. *Let M and M' be even supermatrices of size $(m|n)$ over $\mathcal{R}$, and let A, B, C, D be the blocks of M.*

(a) If M is invertible, then $\operatorname{SDet} M = \det A \, \det(D - CA^{-1}B)^{-1}$.

(b) If M and M' are invertible, then $\operatorname{SDet}(MM') = \operatorname{SDet} M \, \operatorname{SDet} M'$.

(c) For homogeneous supermatrices M and M' of parities α and β, we have $\operatorname{STr}(MM') = (-1)^{\alpha\beta} \operatorname{STr}(M'M)$. *In particular* $\operatorname{STr}(MM') = \operatorname{STr}(M'M)$ *for even M and M'.*

(d) The series $e^M = \sum_{s \geq 0} M^s / s!$ converges, e^M is even and invertible, and $\operatorname{SDet}(e^M) = e^{\operatorname{STr} M}$.

(e) Suppose that the coordinates are split into two groups, each of which may contain even and odd coordinates, and that M is invertible and block upper triangular with respect to this split, with diagonal blocks M_1 and M_2, which are even supermatrices. Then $\operatorname{SDet} M = \operatorname{SDet} M_1 \, \operatorname{SDet} M_2$.

In (e), the rows and columns of M are reordered so that the coordinates of the first group come first, and within each group, even coordinates precede odd ones. Statements (a) and (b) are [132, Corollary 3.6.3 and Theorem 3.6.2], and (c) is stated in [132, Section 3.5]. For block diagonal supermatrices with complex entries, [132, Section 3.6] chooses the definition of the superdeterminant so that (d) holds. For (a)–(e), see also [8].

Proof. We begin with (a). We set $F = A^{-1}B$ and $F' = D^{-1}C$, whose entries are odd. Since $A - BD^{-1}C = A(\mathrm{Id} - FF')$ and $D - CA^{-1}B = D(\mathrm{Id} - F'F)$, it suffices to show that

$$\det(\mathrm{Id} - FF') \, \det(\mathrm{Id} - F'F) = 1. \tag{3.10}$$

For $s \geq 1$, we write $(FF')^s = FZF'$ with $Z = (F'F)^{s-1}$, whose entries are even. Using the fact that the odd entries of F and F' anticommute and commute with the even entries of Z, we find

$$\operatorname{Tr}\big((FF')^s\big) = \sum_{a,b,c} F_{ab} Z_{bc} F'_{ca} = -\sum_{a,b,c} F'_{ca} F_{ab} Z_{bc} = -\operatorname{Tr}\big((F'F)^s\big).$$

The matrices FF' and $F'F$ have nilpotent even entries. The identity

$$\det(\mathrm{Id} - K) = \exp\Big(-\sum_{s\geq 1} \frac{\operatorname{Tr}(K^s)}{s}\Big)$$

holds for complex matrices K near zero, and by Lemma 3.4(c) and (d) it holds for matrices K with nilpotent even entries, for which the sum is finite. Applying it to $K = FF'$ and $K = F'F$ proves (3.10).

We now prove (b). We first record three special cases, in which A', B', C', D' are the blocks of M'. Let P and P' be invertible matrices with even entries of sizes m and n, and let F and F' be matrices with odd entries of sizes $m \times n$ and $n \times m$. Using the definition of SDet in the first two cases and (a) in the third, we find that

$$\begin{aligned}
\operatorname{SDet}\begin{pmatrix} A'+FC' & B'+FD' \\ C' & D' \end{pmatrix} &= \frac{\det(A'+FC'-(B'+FD')D'^{-1}C')}{\det D'} = \operatorname{SDet} M', \\
\operatorname{SDet}\begin{pmatrix} PA' & PB' \\ P'C' & P'D' \end{pmatrix} &= \frac{\det P \, \det(A'-B'D'^{-1}C')}{\det P' \, \det D'} = \frac{\det P}{\det P'} \operatorname{SDet} M', \qquad (3.11) \\
\operatorname{SDet}\begin{pmatrix} A' & B' \\ F'A'+C' & F'B'+D' \end{pmatrix} &= \frac{\det A'}{\det(F'B'+D'-(F'A'+C')A'^{-1}B')} = \operatorname{SDet} M'.
\end{aligned}$$

The three supermatrices on the left are the products $U_F M'$, $\operatorname{diag}(P, P')M'$, and $L_{F'}M'$, where U_F and $L_{F'}$ denote the block unitriangular supermatrices with off-diagonal blocks F and F'. Taking $M' = \mathrm{Id}$, we see that

$$\operatorname{SDet} U_F = \operatorname{SDet} L_{F'} = 1, \qquad \operatorname{SDet}\operatorname{diag}(P, P') = \frac{\det P}{\det P'}.$$

Multiplying out, one checks that $M = U_F \operatorname{diag}(P, P') L_{F'}$ with

$$F = BD^{-1}, \qquad P = A - BD^{-1}C, \qquad P' = D, \qquad F' = D^{-1}C.$$

Applying the three cases of (3.11) in turn, we find $\operatorname{SDet}(MM') = (\det P / \det P') \operatorname{SDet} M'$, and the case $M' = \mathrm{Id}$ shows that $\det P / \det P' = \operatorname{SDet} M$.

Next, we prove (c), using that $\operatorname{Tr}(FF') = (-1)^{\gamma\gamma'} \operatorname{Tr}(F'F)$ for matrices F and F' whose entries are homogeneous of parities γ and γ'. Applying this to the four products in the diagonal blocks of MM', we find that $\operatorname{STr}(MM')$ and $(-1)^{\alpha\beta} \operatorname{STr}(M'M)$ consist of

the same four terms with the same signs. For instance, if M and M' are even, then $\operatorname{Tr}(BC') = -\operatorname{Tr}(C'B)$ and $\operatorname{Tr}(CB') = -\operatorname{Tr}(B'C)$, and

$$\begin{aligned}\operatorname{STr}(MM') &= \operatorname{Tr}(AA') + \operatorname{Tr}(BC') - \operatorname{Tr}(CB') - \operatorname{Tr}(DD') \\ &= \operatorname{Tr}(A'A) - \operatorname{Tr}(C'B) + \operatorname{Tr}(B'C) - \operatorname{Tr}(D'D) = \operatorname{STr}(M'M).\end{aligned}$$

The other parities are treated in the same way.

To prove (d), note that the supermatrices with entries in a Grassmann algebra $\mathcal{R}$ form a finite-dimensional algebra. The operator norm of left multiplication on this algebra is submultiplicative, and the series for e^M converges absolutely in it. (For superfunctions we argue at each point.) The supermatrix e^M is even, and e^{-M} is its inverse. We set $f(t) = \operatorname{SDet}(e^{tM})$ for $t \in \mathbb{R}$. By (b), the definition of SDet, and the expansion $e^{tM} = \mathrm{Id} + tM + O(t^2)$, we have

$$f(t+t') = f(t)f(t'), \qquad f(t) = 1 + t \operatorname{STr} M + O(t^2).$$

Therefore $f'(t) = f(t) \operatorname{STr} M$, the derivative of $e^{-t \operatorname{STr} M} f(t)$ vanishes, and $f(t) = e^{t \operatorname{STr} M}$, where we used that the even part of $\mathcal{R}$ is commutative. Setting $t = 1$ proves (d).

It remains to prove (e). A reordering of the coordinates that keeps even coordinates even is a conjugation by an invertible even supermatrix, which does not change SDet by (b). In the order in which the coordinates of the first group come first, we write

$$M = \operatorname{diag}(M_1, M_2)(\mathrm{Id} + K),$$

where K has nonzero entries only in the rows of the first group and the columns of the second group. Then

$$K^2 = 0, \qquad \mathrm{Id} + K = e^K, \qquad \operatorname{STr} K = 0,$$

and (d) implies that $\operatorname{SDet}(\mathrm{Id} + K) = 1$. In the standard order, the four blocks of $\operatorname{diag}(M_1, M_2)$ are block diagonal, so that $\operatorname{SDet} \operatorname{diag}(M_1, M_2) = \operatorname{SDet} M_1 \operatorname{SDet} M_2$ by the definition of SDet. By (b), this completes the proof. □

We can now compute Gaussian integrals whose quadratic form mixes commuting and anticommuting variables. For an even supermatrix M of size $(m|n)$ with blocks A, B, C, D, we define the even superfunction

$$(\zeta, \chi)^* M (\zeta, \chi) = \zeta^* A \zeta + \zeta^* B \chi + \chi^* C \zeta + \chi^* D \chi$$

on $\mathbb{C}^{m|2n}$, where $\zeta^* B \chi = \sum_{i,b} \bar{\zeta}_i B_{ib} \chi_b$ and $\chi^* C \zeta = \sum_{b,j} \bar{\chi}_b C_{bj} \zeta_j$.

Proposition 3.12. *Let M be an even supermatrix of size $(m|n)$ with blocks A, B, C, D, whose entries lie in the auxiliary generators, and suppose that* $\operatorname{Re} A_\varnothing > 0$. *Then*

$$\int e^{-(\zeta,\chi)^* M (\zeta,\chi)} \, \mathrm{d}\zeta \, \mathrm{d}\chi = \frac{\det(D - CA^{-1}B)}{\det A},$$

which equals $\operatorname{SDet}(M)^{-1}$ *if M is invertible.*

Proof. The matrix A is invertible, since $\det A_\varnothing \neq 0$. Expanding the product, one checks that

$$(\zeta,\chi)^* M(\zeta,\chi) = (\zeta^* + \chi^* C A^{-1})\, A\, (\zeta + A^{-1}B\chi) + \chi^*(D - CA^{-1}B)\chi.$$

The entries of the columns

$$u = A^{-1}B\chi, \qquad v = (\chi^* C A^{-1})^T$$

are even and nilpotent. By Lemma 3.4(c), the exponential of minus the first term is obtained from $e^{-\zeta^* A \zeta}$ by shifting ζ by u and $\bar\zeta$ by v, as in (3.7), and the second term does not involve ζ.

We integrate over ζ first, coefficientwise in the odd coordinates and the auxiliary generators. The coefficients of $e^{-\zeta^* A \zeta}$ are polynomials in ζ and $\bar\zeta$ times $e^{-\zeta^* A_\varnothing \zeta}$, and all their derivatives are integrable. By Lemma 3.6(e), the shifts by u and v do not change the integral over ζ, and Proposition 3.8(a) implies that

$$\int_{\mathbb{C}^m} e^{-(\zeta,\chi)^* M(\zeta,\chi)}\, \mathrm{d}\zeta = \frac{1}{\det A}\, e^{-\chi^*(D - CA^{-1}B)\chi}.$$

Integrating over χ with Proposition 3.8(b) completes the proof, and Lemma 3.11(a) identifies the result with $\mathrm{SDet}(M)^{-1}$ when M is invertible. □

Applied to a supermatrix of the form $M \otimes \mathrm{Id}_N$, acting on N copies of the coordinates, Proposition 3.12 yields the Nth power of the integral for M. This is how the size of a matrix becomes a parameter in Section 3.5.

3.5 An example: the Gaussian unitary ensemble

We now carry out the supersymmetric method in the simplest case, to show the reader the steps that Section 4 carries out for band matrices. In this section, H denotes a Hermitian matrix from the Gaussian unitary ensemble (GUE), with centered complex Gaussian entries such that $\mathbb{E} H_{ij} H_{kl} = \delta_{il}\delta_{jk} N^{-1}$. This is the covariance (1.1) of the band matrix of Definition 1.1, with the variance profile $J = (\mathrm{Id} - W^2\Delta)^{-1}$ replaced by its limit $N^{-1}\mathbf{1}\mathbf{1}^T$ as $W \to \infty$.

Proposition 3.13. *Let $N \geq 1$, let $z \in \mathbb{C}$ with $\operatorname{Im} z > 0$, and let $w \in \mathbb{C}$. Then*

$$\mathbb{E}\frac{\det(w - H)}{\det(z - H)} = \frac{N}{2\pi}\int_{\mathbb{R}^2} e^{-N(x^2+y^2)/2}\, \frac{(w - \mathrm{i}y)^{N-1}}{(z - x)^{N+1}}\big((z - x)(w - \mathrm{i}y) - 1\big)\, \mathrm{d}x\, \mathrm{d}y. \tag{3.12}$$

The proof has four steps. We represent the ratio as a superintegral over N copies of $\mathbb{C}^{1|2}$, compute the average over H, decouple the resulting quartic exponent by Gaussian integrals over the entries of a supermatrix of size $(1|1)$, and integrate out the N copies. The last step produces the Nth power of a superdeterminant, and the dimension of the final integral does not depend on N.

Proof. We begin with the representation of the ratio. For $i, j \in [N]$, we set

$$\Pi_{ij} = \bar{\zeta}_i \zeta_j + \bar{\chi}_i \chi_j,$$

and we will show in Section 3.7 that these bilinears are supersymmetric. Applying Lemma 3.9 with the upper sign and $z_+ = z$, and using $\det(w - H)/\det(z - H) = \det(H - w)/\det(H - z)$, we obtain

$$\frac{\det(w - H)}{\det(z - H)} = \int e^{\mathrm{i}z\zeta^*\zeta + \mathrm{i}w\chi^*\chi - \mathrm{i}\sum_{i,j} H_{ij}\Pi_{ij}} \, \mathrm{d}\zeta \, \mathrm{d}\chi \tag{3.13}$$

on $\mathbb{C}^{N|2N}$. The coefficients of the integrand are polynomials in ζ, $\bar{\zeta}$, and the entries of H, times a factor of modulus $e^{-(\operatorname{Im} z)|\zeta|^2}$, and we may take the expectation inside the integral.

We next compute the average. For a complex matrix $\pi = (\pi_{ij})$, the variable $\sum_{i,j} H_{ij}\pi_{ij}$ is a complex linear combination of jointly Gaussian variables, and

$$\mathbb{E}\Big(\sum_{i,j} H_{ij}\pi_{ij}\Big)^2 = \frac{1}{N}\sum_{i,j} \pi_{ij}\pi_{ji}, \qquad \mathbb{E}e^{-\mathrm{i}\sum_{i,j} H_{ij}\pi_{ij}} = e^{-\frac{1}{2N}\sum_{i,j} \pi_{ij}\pi_{ji}}.$$

Both sides of the second identity are entire in π, and their derivatives may be computed under the expectation. By Lemma 3.4(c), the identity also holds for $\pi = \Pi$. We set

$$\alpha = \zeta^*\chi = \sum_{i=1}^N \bar{\zeta}_i \chi_i, \qquad \beta = \chi^*\zeta = \sum_{i=1}^N \bar{\chi}_i \zeta_i,$$

which are odd. Reordering the odd factors, we find

$$\sum_{i,j} \bar{\zeta}_i \zeta_j \bar{\chi}_j \chi_i = \sum_{i,j} \bar{\chi}_j \zeta_j \bar{\zeta}_i \chi_i = \beta\alpha = -\alpha\beta, \qquad \sum_{i,j} \bar{\chi}_i \chi_j \bar{\chi}_j \chi_i = -(\chi^*\chi)^2.$$

Expanding $\Pi_{ij}\Pi_{ji}$, in which each of the two cross terms gives $-\alpha\beta$ after summation, we find

$$\sum_{i,j} \Pi_{ij}\Pi_{ji} = (\zeta^*\zeta)^2 - 2\alpha\beta - (\chi^*\chi)^2,$$

and therefore

$$\mathbb{E}\frac{\det(w - H)}{\det(z - H)} = \int e^{\mathrm{i}z\zeta^*\zeta + \mathrm{i}w\chi^*\chi} \, e^{-(\zeta^*\zeta)^2/(2N)} \, e^{(\chi^*\chi)^2/(2N)} \, e^{\alpha\beta/N} \, \mathrm{d}\zeta \, \mathrm{d}\chi. \tag{3.14}$$

We now decouple the three quartic factors by Gaussian integrals; this is the Hubbard–Stratonovich transformation. For complex numbers a and b, we have

$$e^{-a^2/(2N)} = \sqrt{\frac{N}{2\pi}} \int_{\mathbb{R}} e^{-Nx^2/2 - \mathrm{i}xa} \, \mathrm{d}x, \qquad e^{b^2/(2N)} = \sqrt{\frac{N}{2\pi}} \int_{\mathbb{R}} e^{-Ny^2/2 + yb} \, \mathrm{d}y,$$

and by Lemma 3.4(c) these identities hold for the even elements $a = \zeta^*\zeta$ and $b = \chi^*\chi$. For the odd factor, we introduce two new odd coordinates ϑ and $\bar\vartheta$, with the Berezin integral $\int \cdot \, \mathrm{d}\vartheta$ given by the coefficient of $\vartheta\bar\vartheta$. Applying Proposition 3.8(c) with $n = 1$, $B = N$, $\theta = \beta$, and $\bar\theta = \alpha$, we obtain

$$e^{\alpha\beta/N} = \frac{1}{N} \int e^{-N\bar\vartheta\vartheta + \bar\vartheta\beta + \alpha\vartheta} \, \mathrm{d}\vartheta .$$

We insert these three identities into (3.14). The coefficients of the resulting integrand are bounded by a polynomial times $e^{-(\operatorname{Im} z)|\zeta|^2 - N(x^2+y^2)/2}$, and we may integrate over ζ and χ first. Since

$$\bar\vartheta\beta = -\sum_{i=1}^{N} \bar\chi_i \bar\vartheta \zeta_i , \qquad \alpha\vartheta = -\sum_{i=1}^{N} \bar\zeta_i \vartheta \chi_i ,$$

we may write the exponent as

$$-N\frac{x^2+y^2}{2} - N\bar\vartheta\vartheta - \sum_{i=1}^{N} (\zeta_i, \chi_i)^* M (\zeta_i, \chi_i), \qquad M = \begin{pmatrix} \mathrm{i}(x-z) & \vartheta \\ \bar\vartheta & -(\mathrm{i}w + y) \end{pmatrix}.$$

The supermatrix M is even, with a BB entry of real part $\operatorname{Im} z > 0$. Proposition 3.12, applied to $M \otimes \mathrm{Id}_N$ on $\mathbb{C}^{N|2N}$, yields

$$\int e^{-\sum_i (\zeta_i,\chi_i)^* M (\zeta_i,\chi_i)} \, \mathrm{d}\zeta \, \mathrm{d}\chi = \Big(\frac{-(\mathrm{i}w + y) - \bar\vartheta\vartheta/(\mathrm{i}(x-z))}{\mathrm{i}(x-z)} \Big)^N = \Big(\frac{w - \mathrm{i}y}{z-x} + \frac{\bar\vartheta\vartheta}{(z-x)^2} \Big)^N .$$

It remains to integrate over ϑ, x, and y. Since $(\bar\vartheta\vartheta)^2 = 0$, we have

$$e^{-N\bar\vartheta\vartheta} \Big(\frac{w-\mathrm{i}y}{z-x} + \frac{\bar\vartheta\vartheta}{(z-x)^2} \Big)^N = \frac{(w-\mathrm{i}y)^N}{(z-x)^N} + N\bar\vartheta\vartheta \, \frac{(w-\mathrm{i}y)^{N-1}}{(z-x)^{N+1}} \big(1 - (z-x)(w-\mathrm{i}y)\big).$$

The first term on the right does not contain ϑ and has Berezin integral zero. Since the Berezin integral over ϑ is the coefficient of $\vartheta\bar\vartheta$, we have

$$\int \bar\vartheta\vartheta \, \mathrm{d}\vartheta = -1 .$$

Collecting the factor $N^{-1} \cdot N/(2\pi)$ from the three Hubbard–Stratonovich identities, we obtain (3.12). □

We checked (3.12) numerically for several values of z and w, against integration over the eigenvalue density of the GUE for $N \le 3$ and against Monte Carlo simulation for $N = 4$, with the supplementary script `grassmann/gue_check.py`. Exact formulas for general correlation functions of ratios of characteristic polynomials, valid for every unitarily invariant ensemble, were obtained by Fyodorov and Strahov [59] with orthogonal polynomials; the point of the example is the method rather than the formula.

Three features of the proof reappear for band matrices. First, the supermatrix M equals $\mathrm{i}(R - \operatorname{diag}(z, w))$, where

$$R = \begin{pmatrix} x & -\mathrm{i}\vartheta \\ -\mathrm{i}\bar{\vartheta} & \mathrm{i}y \end{pmatrix}, \qquad \operatorname{STr} R^2 = x^2 + y^2 + 2\bar{\vartheta}\vartheta.$$

We have $\operatorname{SDet}(cK) = \operatorname{SDet} K$ for every number $c \neq 0$ and every invertible even supermatrix K of size $(1|1)$, and the proof shows that

$$\mathbb{E}\frac{\det(w - H)}{\det(z - H)} = \frac{1}{2\pi} \int e^{-N \operatorname{STr}(R^2)/2} \operatorname{SDet}\big(\operatorname{diag}(z, w) - R\big)^{-N} \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}\vartheta,$$

where $\operatorname{SDet}(K)^{-1}$ stands for $\det(D - CA^{-1}B)/\det A$, as in Proposition 3.12, when the FF block of K is not invertible. The measure $(2\pi)^{-1} e^{-N \operatorname{STr}(R^2)/2}\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}\vartheta$ has total integral one for every N, since the Gaussian factors of the even and odd entries cancel. In Section 4, the Hubbard–Stratonovich field becomes one supermatrix per site, and the fields at different sites are coupled through the inverse of the variance profile J. Second, the FF entry $\mathrm{i}y$ of R runs over the imaginary axis. The FF quartic factor $e^{(\chi^*\chi)^2/(2N)}$ has the sign opposite to the BB factor, and the factor $e^{-N \operatorname{STr}(R^2)/2}$ is integrable in the FF entry on the imaginary axis but not on the real axis. Third, the BB entry x is real because only $\operatorname{Im} z > 0$ appears. When $E + \mathrm{i}\eta$ and $E - \mathrm{i}\eta$ appear together, the bosonic part of the quartic form is not definite, and the BB entries are integrated over the cycle of Section 4.4 instead.

3.6 Change of variables

For even variables, the change of variables formula multiplies the integrand by the Jacobian determinant. Recall from Lemma 3.6(c) that a linear change of odd variables multiplies it by the inverse determinant instead. Berezin's formula combines these rules for changes of coordinates that mix even and odd variables.

Let $U, V \subset \mathbb{R}^m$ be open, and let $y_1, \dots, y_m$ be even and $\xi'_1, \dots, \xi'_r$ odd superfunctions on $U^{m|r}$. For a superfunction $f = \sum_I f_I \xi^I$ on $V^{m|r}$, we write $f(y, \xi') = \sum_I f_I(y) \xi'^I$, where $f_I(y)$ is defined by Definition 3.3 and ξ'^I is the product of the ξ'_a, $a \in I$, in increasing order. When the body map $y_\varnothing$ is a diffeomorphism from U onto V, we regard (y, ξ') as new coordinates, and we call the even supermatrix of size $(m|r)$

$$\mathcal{J} = \begin{pmatrix} (\partial_{x_i} y_j)_{i,j} & (\partial_{x_i} \xi'_b)_{i,b} \\ (\partial_{\xi_a} y_j)_{a,j} & (\partial_{\xi_a} \xi'_b)_{a,b} \end{pmatrix} \tag{3.15}$$

the *Jacobian supermatrix* of the change of coordinates. The rows of $\mathcal{J}$ are indexed by the old coordinates and the columns by the new ones. The following lemma justifies the word "coordinates"; it is the form of the inverse function theorem for superdomains that we need, and a local version for supermanifolds is [132, Theorem 4.4.1].

Lemma 3.14. *Suppose that $y_\emptyset$ is a diffeomorphism from U onto V and that the body of the matrix $(\partial_{\xi_a}\xi'_b)_{a,b}$ is invertible at every point of U. Then $f \mapsto f(y,\xi')$ is an isomorphism of algebras from $C^\infty(V^{m|r})$ onto $C^\infty(U^{m|r})$, which maps even superfunctions to even ones and odd superfunctions to odd ones.*

Proof. We write $\Phi(f) = f(y,\xi')$. By Lemma 3.4(a), Φ is a homomorphism of algebras. It preserves parity, since the y_j are even and the ξ'_b are odd.

Let $\mathcal{F}_s$, for $s \geq 0$, denote the superfunctions, on $U^{m|r}$ or on $V^{m|r}$, whose monomials all have degree at least s. Each ξ'_b lies in $\mathcal{F}_1$, with part of degree one $\sum_a (D_\emptyset)_{ab}\xi_a$, where $D_\emptyset = (\partial_{\xi_a}\xi'_b)_\emptyset$ is the body of the FF block of $\mathcal{J}$. The soul of each y_j lies in $\mathcal{F}_2$. We therefore have, for $|I| = s$,

$$\Phi(f_I\xi^I) = (f_I \circ y_\emptyset)\prod_{b\in I}\Big(\sum_a (D_\emptyset)_{ab}\xi_a\Big) \mod \mathcal{F}_{s+1},$$

where the product is taken in increasing order of b. Thus Φ maps $\mathcal{F}_s$ into $\mathcal{F}_s$, and on the monomials of degree exactly s, modulo $\mathcal{F}_{s+1}$, it acts as the composition with $y_\emptyset$ followed by the linear change of generators with the invertible matrix $D_\emptyset(x)$. This induced map is bijective, with an inverse given by the composition with $y_\emptyset^{-1}$ and the inverse matrix, whose entries are smooth.

If $\Phi(f) = 0$ and $f \neq 0$, let s be the smallest degree of a monomial of f. Then the part of f of degree s is mapped to zero modulo $\mathcal{F}_{s+1}$, which contradicts the bijectivity of the induced map. Hence Φ is injective. To prove surjectivity, let $f' \in C^\infty(U^{m|r})$, and suppose that $f' \in \mathcal{F}_s$. By the bijectivity of the induced map, there exists f_s, a sum of monomials of degree s with coefficients in $C^\infty(V)$, such that $f' - \Phi(f_s) \in \mathcal{F}_{s+1}$. Repeating this with $f' - \Phi(f_s)$ in place of f', we obtain $f = f_s + f_{s+1} + \cdots + f_r$ with $\Phi(f) = f'$, since $\mathcal{F}_{r+1} = \{0\}$. □

By Lemma 3.14, every superfunction on $U^{m|r}$ can be written uniquely as $f(y,\xi')$ with $f \in C^\infty(V^{m|r})$, and Berezin's formula compares the Berezin integrals in the two coordinate systems.

Theorem 3.15. *In this setting, suppose that $y_\emptyset$ is an orientation-preserving diffeomorphism from U onto V and that the body of the matrix $(\partial_{\xi_a}\xi'_b)_{a,b}$ is invertible at every point of U. Then, for every superfunction f on $V^{m|r}$ whose coefficients have compact support in V,*

$$\int_{V^{m|r}} f \,\mathrm{d}y\,\mathrm{d}\xi' = \int_{U^{m|r}} f(y,\xi')\ \mathrm{SDet}\,\mathcal{J}\,\mathrm{d}x\,\mathrm{d}\xi.$$

This is Berezin's change of variables formula; see [8] and, with another arrangement of the matrix of derivatives, [132, Theorem 4.6.1]. With the Berezinian of a change of coordinates written as in [3, Definition 2.5], it is [3, Theorem 2.7], which is credited there to Leites [84]. Conventions for the arrangement of $\mathcal{J}$ and for the order of the

odd coordinates differ between references, and the plain transpose of $\mathcal{J}$, with rows indexed by the new coordinates, gives a different, incorrect, formula. (The supplementary script `grassmann/berezin_change_check.py` checks the formula, in our arrangement, for a nonlinear change of coordinates on $\mathbb{R}^{1|2}$, and shows that the other arrangement fails.) Linear changes of the odd coordinates and translations of the even coordinates by nilpotent elements are the special cases proved in Lemma 3.6. For $y = x$ and $\xi'_b = \sum_a B_{ba}\xi_a$, we have $\operatorname{SDet}\mathcal{J} = (\det B)^{-1}$, and for $y = x + u$ and $\xi' = \xi$, we have $\operatorname{SDet}\mathcal{J} = 1$.

Boundary terms

For our purposes, boundary terms are the most delicate point of the calculus, and they determine the choice of coordinates in Section 5. The hypothesis of compact support in Theorem 3.15 cannot be dropped, even for changes of coordinates that have $\operatorname{SDet}\mathcal{J} = 1$ and map U onto itself. The following example is known as Rudakov's example (see [3, Example 2.8]). Let $U = (0,\infty)$, $r = 2$, and $g \in C_c^\infty(\mathbb{R})$, and consider the coordinates $y = x + \xi_1\xi_2$ and $\xi' = \xi$ on $U^{1|2}$. The Jacobian supermatrix has the blocks 1, 0, $(\xi_2, -\xi_1)^T$, and Id, so that $\operatorname{SDet}\mathcal{J} = 1$. The function $f = g$, a function of the even coordinate alone, has $\int_{U^{1|2}} f \,\mathrm{d}y\,\mathrm{d}\xi' = 0$. However, $f(y,\xi') = g(x) + g'(x)\xi_1\xi_2$, and therefore

$$\int_{U^{1|2}} f(y,\xi') \operatorname{SDet}\mathcal{J}\,\mathrm{d}x\,\mathrm{d}\xi = \int_0^\infty g'(x)\,\mathrm{d}x = -g(0).$$

The two sides differ by a term at the boundary point $x = 0$, unless $g(0) = 0$. The same computation on $U = \mathbb{R}$ gives zero on each side.

The same phenomenon occurs at the singular point of polar coordinates, and there it carries information. Let $g \in C_c^\infty([0,\infty))$, and consider $F = g(\bar\zeta\zeta + \bar\chi\chi)$ on $\mathbb{C}^{1|2}$. By Definition 3.3, we have $F = g(|\zeta|^2) - g'(|\zeta|^2)\chi\bar\chi$, and therefore

$$\int F\,\mathrm{d}\zeta\,\mathrm{d}\chi = -\frac{1}{\pi}\int_{\mathbb{C}} g'(|\zeta|^2)\,\mathrm{d}\operatorname{Re}\zeta\,\mathrm{d}\operatorname{Im}\zeta = -\int_0^\infty g'(t)\,\mathrm{d}t = g(0). \tag{3.16}$$

However, on $(\mathbb{C}\setminus\{0\})^{1|2}$ we may use the new even coordinates

$$\zeta' = \zeta\Big(1 + \frac{\bar\chi\chi}{2|\zeta|^2}\Big), \qquad \bar\zeta' = \bar\zeta\Big(1 + \frac{\bar\chi\chi}{2|\zeta|^2}\Big),$$

which satisfy $\bar\zeta'\zeta' = \bar\zeta\zeta + \bar\chi\chi$, and keep the odd coordinates $\chi, \bar\chi$. In the coordinates $(\zeta', \chi, \bar\chi)$, the function $F = g(\bar\zeta'\zeta')$ is a function of the even coordinates alone, with Berezin integral zero. The odd coordinates are unchanged, so the BF block of the Jacobian supermatrix vanishes, its FF block is the identity, and $\operatorname{SDet}\mathcal{J}$ is the determinant of its BB block. In real coordinates, $\zeta \mapsto \zeta'$ is a radial dilation by the factor $1 + c/(2|\zeta|^2)$ with $c = \bar\chi\chi$, whose Jacobian determinant $1 - c^2/(4|\zeta|^4)$ reduces to one, since $c^2 = 0$. Therefore $\operatorname{SDet}\mathcal{J} = 1$. The change of coordinates is singular at $\zeta = 0$, where F does not vanish,

and the value $g(0)$ in (3.16) is a boundary term concentrated there. A computation of this kind in polar coordinates on a superdomain of dimension $(2|2)$ is [3, Examples 4.7(ii)].

In the supersymmetric method, one often parametrizes supermatrices by eigenvalues and angles, and the boundary contributions of such changes of variables are known in physics as Efetov–Wegner terms (see [76]). A general formula for them was given by Rothstein [110], and a coordinate-free form by Alldridge, Hilgert, and Palzer [3]. Zirnbauer [139, Section II.A] explains why they matter for the integrals of the supersymmetric method, whose integrands are not compactly supported. Berezin's formula is an identity for compactly supported integrands, and an integrand that does not vanish at the boundary of the body domain, or at a singular point of the coordinates, produces additional terms. In this paper, superintegrals over the symmetric superspace of Section 5 are written in global coordinates that are nonsingular everywhere (Section 5.2). The body has no boundary, and the integrands have integrable coefficients, so that integration by parts produces no boundary terms. At rank one, Section 6 also uses coordinates that are singular at a point, and there the contribution of the singular point is accounted for separately; the example (3.16) shows that localization can identify it.

3.7 Supersymmetry and localization

Recall from (3.8), with $w = z_\pm$, that

$$\int e^{\mp \mathrm{i}[\zeta^*(H-z_\pm)\zeta + \chi^*(H-z_\pm)\chi]}\,\mathrm{d}\zeta\,\mathrm{d}\chi = 1$$

for every Hermitian matrix H. This is forced by a symmetry that exchanges commuting and anticommuting variables. We now explain the mechanism, which also normalizes the sigma model in Section 5. We work on the flat superspace $\mathbb{C}^{n|2n}$ of Definition 3.7.

Definition 3.16. We define the odd vector fields

$$\mathfrak{q}_+ = \sum_{j=1}^{n} \left(\chi_j \partial_{\zeta_j} - \bar{\zeta}_j \partial_{\bar{\chi}_j}\right), \qquad \mathfrak{q}_- = \sum_{j=1}^{n} \left(\bar{\chi}_j \partial_{\bar{\zeta}_j} + \zeta_j \partial_{\chi_j}\right)$$

on $\mathbb{C}^{n|2n}$, and we define the *supercharge* to be $\mathfrak{q} = \mathfrak{q}_+ + \mathfrak{q}_-$. We call a superfunction F *supersymmetric*, or $\mathfrak{q}$-*closed*, if $\mathfrak{q}F = 0$, and $\mathfrak{q}_\pm$-*closed* if $\mathfrak{q}_\pm F = 0$.

We have $\mathfrak{q}\zeta_j = \chi_j$, $\mathfrak{q}\bar{\zeta}_j = \bar{\chi}_j$, $\mathfrak{q}\chi_j = \zeta_j$, and $\mathfrak{q}\bar{\chi}_j = -\bar{\zeta}_j$ on the coordinates; the field $\mathfrak{q}_+$ acts only on ζ_j and $\bar{\chi}_j$, and $\mathfrak{q}_-$ only on $\bar{\zeta}_j$ and χ_j. Squares of odd vector fields are even vector fields, and computing them on the coordinates, one finds that $\mathfrak{q}_+^2 = \mathfrak{q}_-^2 = 0$ and

$$\mathfrak{q}^2 = \mathfrak{q}_+\mathfrak{q}_- + \mathfrak{q}_-\mathfrak{q}_+ = \sum_{j=1}^{n} \left(\zeta_j \partial_{\zeta_j} + \chi_j \partial_{\chi_j} - \bar{\zeta}_j \partial_{\bar{\zeta}_j} - \bar{\chi}_j \partial_{\bar{\chi}_j}\right).$$

The operator $\mathfrak{i}\mathfrak{q}^2$ generates the phase rotations, which multiply ζ and χ by $e^{\mathrm{i}t}$ and $\bar\zeta$ and $\bar\chi$ by $e^{-\mathrm{i}t}$, for $t \in \mathbb{R}$. Every supersymmetric function F is invariant under these rotations, since $\mathfrak{q}^2 F = \mathfrak{q}(\mathfrak{q}F) = 0$. Thus $\mathfrak{q}_+$ and $\mathfrak{q}_-$ have square zero, and so has $\mathfrak{q}$ on invariant functions. Sign conventions for these fields differ between references; ours is adapted to the quadratic forms $\zeta^* A\zeta + \chi^* A\chi$ of Lemma 3.9.

Lemma 3.17. *Let $i, j \in [n]$, and set*

$$\Pi_{ij} = \bar\zeta_i \zeta_j + \bar\chi_i \chi_j, \qquad \sigma = \sum_{l=1}^n \bar\zeta_l \chi_l, \qquad \sigma_+ = -\sum_{l=1}^n \bar\chi_l \zeta_l$$

on $\mathbb{C}^{n|2n}$, and let $\mathfrak{q}'$ be one of the fields $\mathfrak{q}$, $\mathfrak{q}_+$, and $\mathfrak{q}_-$.

(a) We have $\mathfrak{q}'\Pi_{ij} = 0$, and $\mathfrak{q}'\, g(\Pi) = 0$ for every entire function g of the n^2 entries of the matrix $\Pi = (\Pi_{ij})$.

(b) We have $\mathfrak{q}\sigma = \mathfrak{q}_-\sigma = \mathfrak{q}_+\sigma_+ = \zeta^*\zeta + \chi^*\chi$.

(c) For every superfunction F on $\mathbb{C}^{n|2n}$ of class C^1 such that F, its first-order partial derivatives, and $\mathfrak{q}'F$ are integrable, we have $\int \mathfrak{q}'F \,\mathrm{d}\zeta\,\mathrm{d}\chi = 0$.

Proof. Using that $\mathfrak{q}_\pm$ are odd derivations, we compute

$$\begin{aligned}
\mathfrak{q}_+(\bar\zeta_i\zeta_j) &= \bar\zeta_i\chi_j, \quad \mathfrak{q}_+(\bar\chi_i\chi_j) = (\mathfrak{q}_+\bar\chi_i)\chi_j - \bar\chi_i\,\mathfrak{q}_+\chi_j = -\bar\zeta_i\chi_j,\\
\mathfrak{q}_-(\bar\zeta_i\zeta_j) &= \bar\chi_i\zeta_j, \quad \mathfrak{q}_-(\bar\chi_i\chi_j) = (\mathfrak{q}_-\bar\chi_i)\chi_j - \bar\chi_i\,\mathfrak{q}_-\chi_j = -\bar\chi_i\zeta_j,
\end{aligned}$$

so that

$$\mathfrak{q}_+\Pi_{ij} = \mathfrak{q}_-\Pi_{ij} = \mathfrak{q}\Pi_{ij} = 0.$$

The second claim of (a) follows from Lemma 3.4(b). Next, we have

$$\mathfrak{q}_+(\bar\zeta_l\chi_l) = 0, \qquad \mathfrak{q}_-(\bar\zeta_l\chi_l) = \bar\chi_l\chi_l + \bar\zeta_l\zeta_l, \qquad \mathfrak{q}_+(-\bar\chi_l\zeta_l) = \bar\zeta_l\zeta_l + \bar\chi_l\chi_l,$$

and summing over l gives (b).

To prove (c), note first that the assumed integrability of $\mathfrak{q}'F$ makes its Berezin integral well defined. The top coefficient of $\zeta_l\partial_{\chi_l}F$ vanishes, since this superfunction does not contain χ_l; the same holds for $\bar\zeta_l\partial_{\bar\chi_l}F$. Up to sign, the top coefficient of $\chi_l\partial_{\zeta_l}F$ is the derivative in ζ_l of a coefficient of F, whose integral vanishes by Lemma 3.6(b), and the same holds for $\bar\chi_l\partial_{\bar\zeta_l}F$. Each of the three fields is a combination of these four kinds of terms. □

Statement (c) says that $\mathfrak{q}$, $\mathfrak{q}_+$, and $\mathfrak{q}_-$ have zero divergence for the flat Berezin integral, and statements (a) and (b) say that $\zeta^*\zeta + \chi^*\chi$ is closed and exact for each of them. The deformation step of the localization argument uses only these three properties, and we state it for a general odd vector field, which we also denote by $\mathfrak{q}$, and for a Berezin integral

with a density, so that Section 5 can apply it on a curved superspace. For such a field, $\mathfrak{q}^2$ is in general a nonzero even vector field, and the lemma assumes only that $\mathfrak{q}(\mathfrak{q}\sigma) = 0$ for the exponent $\mathfrak{q}\sigma$, which would be automatic if $\mathfrak{q}^2 = 0$. For an even superfunction ρ on $U^{m|r}$ and a superfunction F such that ρF is integrable, we write $\int_\rho F = \int_{U^{m|r}} \rho F \, \mathrm{d}x \, \mathrm{d}\xi$.

Lemma 3.18. *Let $\mathfrak{q}$ be an odd vector field on $U^{m|r}$, let ρ be an even superfunction, and let F and σ be superfunctions of class C^1, with σ odd. Suppose that $\mathfrak{q}F = 0$ and $\mathfrak{q}(\mathfrak{q}\sigma) = 0$. Let $\mathcal{T} \subset \mathbb{R}$ be an open interval, and suppose that the following conditions hold for all $t \in \mathcal{T}$.*

(i) The superfunctions $\rho F e^{-t\mathfrak{q}\sigma}$ and $\rho(\mathfrak{q}\sigma)Fe^{-t\mathfrak{q}\sigma}$ are integrable, and the coefficients of the second are bounded, locally uniformly in t, by an integrable function.

(ii) We have $\int_\rho \mathfrak{q}(\sigma F e^{-t\mathfrak{q}\sigma}) = 0$.

Then $t \mapsto \int_\rho F e^{-t\mathfrak{q}\sigma}$ is constant on $\mathcal{T}$.

Proof. The superfunction $\mathfrak{q}\sigma$ is even, and Lemma 3.4(b) yields

$$\mathfrak{q} e^{-t\mathfrak{q}\sigma} = -t\, e^{-t\mathfrak{q}\sigma}\, \mathfrak{q}(\mathfrak{q}\sigma) = 0,$$

where we used the assumption $\mathfrak{q}(\mathfrak{q}\sigma) = 0$ in the last step. Since $\mathfrak{q}$ is an odd derivation and σ is odd, we have

$$\mathfrak{q}(\sigma F e^{-t\mathfrak{q}\sigma}) = (\mathfrak{q}\sigma) F e^{-t\mathfrak{q}\sigma} - \sigma\, \mathfrak{q}(F e^{-t\mathfrak{q}\sigma}) = (\mathfrak{q}\sigma) F e^{-t\mathfrak{q}\sigma},$$

where we used $\mathfrak{q}F = 0$. By (i), we may differentiate under the integral, and we find

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\rho F e^{-t\mathfrak{q}\sigma} = -\int_\rho (\mathfrak{q}\sigma) F e^{-t\mathfrak{q}\sigma} = -\int_\rho \mathfrak{q}(\sigma F e^{-t\mathfrak{q}\sigma}) = 0,$$

where the second equality follows from the previous display and the third from (ii). □

In words, a factor $e^{-t\mathfrak{q}\sigma}$ with a $\mathfrak{q}$-exact and $\mathfrak{q}$-closed exponent can be inserted into the integral of a $\mathfrak{q}$-closed function without changing it. As $t \to \infty$, this factor concentrates the integral near the zeros of the body of $\mathfrak{q}\sigma$, where it can be computed. Schwarz and Zaboronsky [113, Theorem 1] prove a general localization theorem of this kind on compact supermanifolds, and they explain that the dimensional reduction of Parisi and Sourlas [103] can be obtained from such theorems. The superspaces of this paper are not compact, and the decay of the integrands replaces compactness. In flat space, the localization theorem reads as follows. For the field $\mathfrak{q}$, it is [7, Theorem 11.4.5], where it is formulated for differential forms, and the same argument on the hyperbolic superplane is [28, Appendix C, Lemma 16]. We include the short proof, since it also covers $\mathfrak{q}_\pm$ and Section 5 repeats its steps in a curved setting.

Proposition 3.19. *Let $\mathfrak{q}'$ be one of the fields $\mathfrak{q}$, $\mathfrak{q}_+$, and $\mathfrak{q}_-$. Let F be a superfunction on $\mathbb{C}^{n|2n}$ of class C^1 such that F and its first-order partial derivatives are integrable, and suppose that $\mathfrak{q}'F = 0$. Then*

$$\int F \, \mathrm{d}\zeta \, \mathrm{d}\chi = F_\emptyset(0).$$

Proof. Let $\sigma' = \sigma_+$ if $\mathfrak{q}' = \mathfrak{q}_+$, and $\sigma' = \sigma$ otherwise, with σ and σ_+ as in Lemma 3.17. By Lemma 3.17(b), we have

$$\mathfrak{q}'\sigma' = \zeta^*\zeta + \chi^*\chi = \sum_{j=1}^n \Pi_{jj},$$

and Lemma 3.17(a) therefore implies that $\mathfrak{q}'(\mathfrak{q}'\sigma') = 0$. We first apply Lemma 3.18 with $\mathfrak{q}'$ and σ' in place of $\mathfrak{q}$ and σ, $\rho = 1$, and $\mathcal{T} = (0, \infty)$. The $\chi_j\bar{\chi}_j$ are even and $(\chi_j\bar{\chi}_j)^2 = 0$, so that

$$e^{-t\mathfrak{q}'\sigma'} = e^{-t|\zeta|^2} \prod_{j=1}^n (1 + t\chi_j\bar{\chi}_j). \tag{3.17}$$

Since $|\zeta|^p e^{-t|\zeta|^2}$ is bounded for $p \le 3$, the coefficients of $Fe^{-t\mathfrak{q}'\sigma'}$, of $(\mathfrak{q}'\sigma')Fe^{-t\mathfrak{q}'\sigma'}$, and of $\sigma'Fe^{-t\mathfrak{q}'\sigma'}$, and the first derivatives of the last, are bounded by a constant times sums of the moduli of the coefficients of F and of their first derivatives, with a constant that is locally uniform in $t > 0$. This verifies (i). Moreover, the identity

$$\mathfrak{q}'\big(\sigma'Fe^{-t\mathfrak{q}'\sigma'}\big) = (\mathfrak{q}'\sigma')Fe^{-t\mathfrak{q}'\sigma'}$$

from the proof of Lemma 3.18 shows that this $\mathfrak{q}'$-derivative is integrable as well. Thus Lemma 3.17(c), applied to $\sigma'Fe^{-t\mathfrak{q}'\sigma'}$, verifies (ii). Therefore $\int Fe^{-t\mathfrak{q}'\sigma'} \, \mathrm{d}\zeta \, \mathrm{d}\chi$ is constant in $t > 0$.

For $K \subset [n]$, let $F_{(K)}$ denote the coefficient of $\prod_{j\in K} \chi_j\bar{\chi}_j$ in F, so that $F_{(\emptyset)} = F_\emptyset$ and $F_{([n])} = F_{\mathrm{top}}$. By (3.17), the top coefficient of $Fe^{-t\mathfrak{q}'\sigma'}$ is

$$(Fe^{-t\mathfrak{q}'\sigma'})_{\mathrm{top}} = e^{-t|\zeta|^2} \sum_{K\subset[n]} t^{|K|} F_{([n]\setminus K)}. \tag{3.18}$$

We now let $t \downarrow 0$. The right side of (3.18) converges pointwise to F_{top}, and for $t \le 1$ it is bounded by $\sum_K |F_{(K)}|$, which is integrable. By dominated convergence, we have

$$\lim_{t\downarrow 0} \int Fe^{-t\mathfrak{q}'\sigma'} \, \mathrm{d}\zeta \, \mathrm{d}\chi = \int F \, \mathrm{d}\zeta \, \mathrm{d}\chi.$$

Next, we let $t \to \infty$. For every integrable function $\tilde{g}$ on $\mathbb{C}^n$ that is continuous at zero, we have

$$\lim_{t\to\infty} t^n \int_{\mathbb{C}^n} e^{-t|\zeta|^2} \tilde{g}(\zeta) \, \mathrm{d}\zeta = \tilde{g}(0),$$

where we used that $t^n e^{-t|\zeta|^2} \le t^n e^{-t}$ for $|\zeta| \ge 1$ and that

$$t^n \int_{\mathbb{C}^n} e^{-t|\zeta|^2} \, \mathrm{d}\zeta = 1.$$

We apply this to the terms of (3.18). The term with $K = [n]$ converges to $F_\varnothing(0)$, while each term with $|K| < n$ carries a factor $t^{|K|-n}$ in front of a convergent quantity and tends to zero. Therefore

$$\lim_{t\to\infty} \int F e^{-t\mathfrak{q}'\sigma'} \, \mathrm{d}\zeta \, \mathrm{d}\chi = F_\varnothing(0).$$

Since the integral is constant in t, this completes the proof. □

(For $n = 1$ and $F = g(\bar\zeta\zeta + \bar\chi\chi)$ with $g \in C_c^\infty([0,\infty))$, the function F is supersymmetric by Lemma 3.4(b) and Lemma 3.17(a), and Proposition 3.19 gives the value $g(0)$ found in (3.16).)

The proof uses three properties: the divergence identity of Lemma 3.17(c), the exactness of the localizing exponent, and the invariance of the Berezin integral under the rescaling $\zeta \mapsto t^{-1/2}\zeta$, $\chi \mapsto t^{-1/2}\chi$, $\bar\chi \mapsto t^{-1/2}\bar\chi$, whose Berezinian is one. The last property appears in the proof as the cancellation of $t^{|K|}$ against t^{-n} in the term with $K = [n]$. The odd vector fields generated by the odd symmetries of the sigma model on the symmetric superspace of Section 5 have the same properties, and the proof of the localization lemma Lemma 5.19 follows the same three steps.

In applications, the integrand is often a polynomially bounded function times a supersymmetric Gaussian. The following consequence of Proposition 3.19 treats this case; it is the flat-space prototype of the zero-source normalization Corollary 5.20. For $A \in \mathbb{C}^{n\times n}$ with $\operatorname{Re} A > 0$, we let $\lambda_A > 0$ denote the smallest eigenvalue of $\operatorname{Re} A$.

Corollary 3.20. *Let $A \in \mathbb{C}^{n\times n}$ with $\operatorname{Re} A > 0$, and let $\mathfrak{q}'$ be one of the fields $\mathfrak{q}$, $\mathfrak{q}_+$, and $\mathfrak{q}_-$. Let f be a superfunction on $\mathbb{C}^{n|2n}$ of class C^1 such that $\mathfrak{q}' f = 0$, and suppose that there exist $C > 0$ and $\gamma < \lambda_A$ such that the coefficients of f and their first-order partial derivatives are bounded by $Ce^{\gamma|\zeta|^2}$. Then*

$$\int f \, e^{-\zeta^* A\zeta - \chi^* A\chi} \, \mathrm{d}\zeta \, \mathrm{d}\chi = f_\varnothing(0).$$

In particular, $\int P(\Pi)\, e^{-\zeta^* A\zeta - \chi^* A\chi} \, \mathrm{d}\zeta \, \mathrm{d}\chi = P(0)$ for every polynomial P in the entries of Π, and $\int e^{-\zeta^* A\zeta - \chi^* A\chi} \, \mathrm{d}\zeta \, \mathrm{d}\chi = 1$.

Proof. Observe that

$$\zeta^* A\zeta + \chi^* A\chi = \sum_{i,j=1}^{n} A_{ij}\Pi_{ij},$$

so $e^{-\zeta^* A\zeta - \chi^* A\chi}$ is an entire function of Π, which is $\mathfrak{q}'$-closed by Lemma 3.17(a). Since $\mathfrak{q}' f = 0$, the product $F = f e^{-\zeta^* A\zeta - \chi^* A\chi}$ satisfies $\mathfrak{q}' F = 0$. The coefficients of $e^{-\chi^* A\chi}$ are constants, and

$$|e^{-\zeta^* A\zeta}| = e^{-\zeta^* (\operatorname{Re} A)\zeta} \le e^{-\lambda_A |\zeta|^2}.$$

The coefficients of F and their first derivatives are therefore bounded by a constant times $(1+|\zeta|)e^{-(\lambda_A-\gamma)|\zeta|^2}$, which is integrable. Then Proposition 3.19 implies the first claim, since $F_\varnothing(0) = f_\varnothing(0)$. By Lemma 3.17(a), the polynomial $P(\Pi)$ is closed for each of the three fields. The coefficients of $P(\Pi)$ are polynomials in ζ and $\bar\zeta$, and $P(\Pi)_\varnothing(0) = P(0)$. □

For $f = 1$, the corollary also follows from Proposition 3.8(d). For $f = \Pi_{ji}$, it says that the two insertions of Proposition 3.8(d) have opposite contributions, since $\bar\zeta_j\zeta_i = \zeta_i\bar\zeta_j$ and $\bar\chi_j\chi_i = -\chi_i\bar\chi_j$. By Lemma 3.6(c), a linear change of the odd coordinates multiplies the Berezin integral by a constant, so the corollary also identifies normalized Gaussian integrals in other coordinates. For instance, in the real coordinates $\zeta_j = x_j + \mathrm{i}y_j$ and the odd coordinates $\alpha_j = \bar\chi_j$ and $\beta_j = \chi_j/2$, we have

$$\zeta^*\zeta + \chi^*\chi = \sum_{j=1}^n \big(x_j^2 + y_j^2 + 2\alpha_j\beta_j\big), \qquad \mathfrak{q}_+ = \sum_{j=1}^n \big(\beta_j(\partial_{x_j} - \mathrm{i}\partial_{y_j}) - (x_j - \mathrm{i}y_j)\partial_{\alpha_j}\big).$$

With $A = \frac{1}{2}\mathrm{Id}$, or with a diagonal A with $\operatorname{Re} A > 0$, the corollary then says that the normalized Gaussian integral of a polynomially bounded $\mathfrak{q}_+$-closed function is the value of its body at the origin. This is the form in which localization is used for the normal oscillator in Section 16.

SECTION 4

The supersymmetric representation of the band matrix

In this section, we write the source ratios $T_N^{(k)}$ of (1.4) as the integral of a one-dimensional chain. The chain has one supermatrix-valued variable at each site of $[N]$, a weight attached to each site, and a Gaussian link between neighboring sites. By Proposition 4.11, the chain integral is a bilinear pairing in which a power of an explicit transfer operator T_Γ appears. In Part IV, we start from this representation and from the identity for partial chains in Proposition 4.13, and we prove that the chain converges to the sigma model of Section 5. In Part III, we use Lemma 4.14, which relates resolvent entries of H to derivatives of source ratios and to insertions in the chain.

The derivation is the supersymmetric method, which Section 3.5 carries out for the GUE, applied to the band matrix of Definition 1.1. The supersymmetric approach to band matrices goes back to Fyodorov and Mirlin [53], who used it to derive, at the level of physics, the one-dimensional sigma model. Our derivation has three steps. We write each determinant in $T_N^{(k)}$ as a Gaussian integral, bosonic for the inverse determinants and fermionic for the others, we average over H, and we decouple the resulting quartic interaction by a Hubbard–Stratonovich transformation, with one supermatrix for each site. For the density of states, which involves only the upper half plane, these steps are carried out for the same variance profile in [29, Proposition 2.1]. For block band matrices, M. Shcherbina and T. Shcherbina derived an integral representation of the two-point function, which involves both half planes, with the noncompact hyperbolic sector and the anticommuting variables [117, Proposition 2.1] (see also [119]).

Two points need care. The first is the choice of the variance profile, which makes the decoupled integrand a chain with nearest-neighbor coupling (see Section 4.3). The second is that $T_N^{(k)}$ contains resolvents in the upper and the lower half plane. In this case, the commuting variables of the Hubbard–Stratonovich transformation cannot be integrated over Hermitian matrices. In [117], the commuting variables of each block are instead integrated out by a bosonization formula ([117, Proposition 7.1], taken there from [51]). Such a formula replaces an integral over several commuting vectors, of a function of the matrix of their inner products, by an integral over this matrix. It applies there because each block carries as many vector variables as its size, which is at least two. In our model, each site carries a single vector variable, and we integrate the commuting

variables of the Hubbard–Stratonovich transformation over a cycle introduced by Schäfer and Wegner [111]. We prove the identity that justifies this in Appendix B. After the exact representation, Section 4.6 describes in elementary terms the saddle manifold, near which Part IV analyzes the chain. It is a union of $2k+1$ conjugation orbits, each consisting of critical points of the onsite weight. In Section 5, we describe its balanced component as a homogeneous superspace.

We use the Grassmann algebras, supermatrices, the supertrace STr, the superdeterminant SDet, and the Berezin integral of Section 3. Throughout this section, the integers $k \geq 1$ and $N \geq 2$, the parameters $\varepsilon > 0$ and $W > 0$, and the energy $E \in (-2, 2)$ are fixed. Recall s_E, $\rho(E)$, and a_N from (1.3), and recall that $\kappa_E = E/2$, so that $\kappa_E^2 + s_E^2 = 1$. We let $\|A\|$ denote the Hilbert–Schmidt norm of a complex square matrix A and $\|A\|_{\mathrm{op}}$ its operator norm, and we set $\operatorname{Re} A = (A + A^*)/2$ and $\operatorname{Im} A = (A - A^*)/(2\mathrm{i})$. We write $A > 0$ if A is Hermitian and positive definite.

4.1 Source ratios as superdeterminants

We first collect the spectral parameters of $T_N^{(k)}$ in one diagonal supermatrix. The supermatrices of this section have size $(2k|2k)$. We index the rows of the BB block, and those of the FF block, by $2k$ *flavors*, and we call the flavors $1, \dots, k$ *retarded* and the flavors $k+1, \dots, 2k$ *advanced*. We let $\operatorname{diag}(A \mid D)$ denote the supermatrix with BB block A, FF block D, and vanishing odd blocks.

Definition 4.1. We set $\Lambda_{\mathrm{B}} = \Lambda_{\mathrm{F}} = \operatorname{diag}(\mathrm{Id}_k, -\mathrm{Id}_k)$ and $\Lambda = \operatorname{diag}(\Lambda_{\mathrm{B}} \mid \Lambda_{\mathrm{F}})$. We set $\eta_0 = a_N^{-1}$ and $\eta = \varepsilon\eta_0$. For $\Gamma_{\mathrm{F}} \in \mathbb{C}^{2k \times 2k}$, we call

$$\Gamma = \operatorname{diag}(\varepsilon\Lambda_{\mathrm{B}} \mid \Gamma_{\mathrm{F}})$$

a *source supermatrix*. For $u \in \mathbb{R}^k$ and $b \in (0, \infty)^k$, we let $\Gamma(u, b)$ denote the source supermatrix with

$$\Gamma_{\mathrm{F}} = \operatorname{diag}(b_1 - \mathrm{i}u_1, \dots, b_k - \mathrm{i}u_k, -b_1 - \mathrm{i}u_1, \dots, -b_k - \mathrm{i}u_k),$$

and we set $\Gamma_0 = \Gamma(0, \varepsilon\mathbf{1}) = \varepsilon\Lambda$.

The diagonal entries of $E + \mathrm{i}\eta_0\Gamma(u, b)$ are the spectral parameters of the determinants in $T_N^{(k)}$. Recall from Section 1.2 that $z_a = E + (u_a + \mathrm{i}b_a)/a_N$ for $a \in [k]$ and $z_0 = E + \mathrm{i}\varepsilon/a_N$. Then

$$E + \mathrm{i}\eta_0\Gamma(u, b) = \operatorname{diag}\left(z_0\mathrm{Id}_k, \bar{z}_0\mathrm{Id}_k \mid z_1, \dots, z_k, \bar{z}_1, \dots, \bar{z}_k\right). \tag{4.1}$$

We will also perturb the spectral parameter at single sites, which produces eigenvector observables (see Lemma 4.14). We call a family $Y = (Y_i)_{i \in [N]}$ of supermatrices $Y_i = \operatorname{diag}(Y_i^{\mathrm{B}} \mid Y_i^{\mathrm{F}})$ with complex $2k \times 2k$ blocks a family of *local sources*, and we set $\|Y^{\mathrm{B}}\|_{\mathrm{op}} =$

$\max_i \|Y_i^{\mathrm{B}}\|_{\mathrm{op}}$. Given a family $(A_i)_{i\in[N]}$ of $(2k|2k)$ supermatrices, we let $A - H$ denote the supermatrix of size $(2kN|2kN)$ with blocks

$$(A-H)_{ij} = \delta_{ij}A_i - H_{ij}\mathrm{Id}, \qquad i,j \in [N],$$

where Id is the identity supermatrix of size $(2k|2k)$. We write $A - H$ as well when $A_i = A$ for all i.

Definition 4.2. For a source supermatrix Γ and local sources Y with $\|Y^{\mathrm{B}}\|_{\mathrm{op}} < \eta$, we define

$$T_N^{(k)}(\Gamma; Y) = \mathbb{E}\,\mathrm{SDet}\left(E + \mathrm{i}\eta_0\Gamma + Y - H\right)^{-1}, \tag{4.2}$$

and we set $T_N^{(k)}(\Gamma) = T_N^{(k)}(\Gamma; 0)$.

We call the sites $i_1 < \cdots < i_m$ of Definition 7.5 the *marked sites*. When $Y_i = 0$ outside the marked sites, $T_N^{(k)}(\Gamma(u,b); Y)$ is the function defined there, through which Part III uses local sources. Since the supermatrix in (4.2) has vanishing odd blocks, its superdeterminant is the ratio of the determinants of its BB and FF blocks.

Lemma 4.3. *Let Γ be a source supermatrix and Y local sources with $\|Y^{\mathrm{B}}\|_{\mathrm{op}} < \eta$. Then the BB block of $E + \mathrm{i}\eta_0\Gamma + Y - H$ is invertible for every Hermitian H, and $T_N^{(k)}(\Gamma; Y)$ is finite and holomorphic in Γ_{F} and in the entries of Y. Moreover, for all $u \in \mathbb{R}^k$ and $b \in (0,\infty)^k$,*

$$T_N^{(k)}(\Gamma(u,b)) = T_N^{(k)}(u,b), \qquad T_N^{(k)}(\Gamma_0) = 1, \qquad \mathrm{STr}\,\Gamma_0^2 = 0. \tag{4.3}$$

Proof. Let D_{B} denote the BB block of $E + \mathrm{i}\eta_0\Gamma + Y - H$, which has the blocks

$$(D_{\mathrm{B}})_{ij} = \delta_{ij}\big(E + \mathrm{i}\eta\Lambda_{\mathrm{B}} + Y_i^{\mathrm{B}}\big) - H_{ij}\mathrm{Id}_{2k}, \qquad i,j \in [N].$$

Since the matrices $E\Lambda_{\mathrm{B}}$ and $H \otimes \Lambda_{\mathrm{B}}$ are Hermitian, we have

$$\mathrm{Im}\left((\mathrm{Id}_N \otimes \Lambda_{\mathrm{B}})D_{\mathrm{B}}\right) = \eta\mathrm{Id} + \mathrm{diag}_{i\in[N]}\left(\mathrm{Im}(\Lambda_{\mathrm{B}}Y_i^{\mathrm{B}})\right) \geq (\eta - \|Y^{\mathrm{B}}\|_{\mathrm{op}})\mathrm{Id} > 0.$$

Recall that a matrix A with $\mathrm{Im}\,A > 0$ is invertible and satisfies $|\det A| \geq \det \mathrm{Im}\,A$. Indeed, the matrix $K = (\mathrm{Im}\,A)^{-1/2}(\mathrm{Re}\,A)(\mathrm{Im}\,A)^{-1/2}$ is Hermitian, and $A = (\mathrm{Im}\,A)^{1/2}(K + \mathrm{i}\mathrm{Id})(\mathrm{Im}\,A)^{1/2}$. Taking determinants, we find

$$|\det A| = \det(\mathrm{Im}\,A)\prod_{\mu}|\mu + \mathrm{i}| \geq \det \mathrm{Im}\,A,$$

where μ runs over the eigenvalues of K, which are real. Together with the previous display, this inequality implies that

$$|\det D_{\mathrm{B}}|^{-1} \leq (\eta - \|Y^{\mathrm{B}}\|_{\mathrm{op}})^{-2kN}$$

for every Hermitian H. The determinant of the FF block is a polynomial in the entries of H, Γ_{F}, and Y^{F}, and the entries of H have Gaussian tails. Dominated convergence now shows that the expectation is finite and holomorphic.

We now prove (4.3). The two blocks of $E + \mathrm{i}\eta_0\Gamma(u,b) - H$ are block diagonal over the flavors by (4.1), so that

$$\operatorname{SDet}\big(E + \mathrm{i}\eta_0\Gamma(u,b) - H\big)^{-1} = \prod_{a=1}^{k} \frac{\det(z_a - H)\det(\bar{z}_a - H)}{\det(z_0 - H)\det(\bar{z}_0 - H)} = \prod_{\alpha=1}^{N}\prod_{a=1}^{k} \frac{(x_\alpha - u_a)^2 + b_a^2}{x_\alpha^2 + \varepsilon^2}.$$

In the second step, we used that for $x = a_N(\lambda - E)$,

$$(z_a - \lambda)(\bar{z}_a - \lambda) = \frac{(x - u_a)^2 + b_a^2}{a_N^2}, \qquad (z_0 - \lambda)(\bar{z}_0 - \lambda) = \frac{x^2 + \varepsilon^2}{a_N^2}.$$

Taking expectations yields the first identity. For $\Gamma_0 = \varepsilon\Lambda$, the BB and FF blocks coincide, which implies that $T_N^{(k)}(\Gamma_0) = 1$, and we have

$$\operatorname{STr}\Gamma_0^2 = \varepsilon^2(\operatorname{Tr}\mathrm{Id}_{2k} - \operatorname{Tr}\mathrm{Id}_{2k}) = 0. \qquad \square$$

By Theorem 1.6, the limits of the source ratios $T_N^{(k)}(u,b)$ determine the limit of the point process Ξ_N, and the local sources give access to eigenvector observables. The rest of the section rewrites $T_N^{(k)}(\Gamma; Y)$.

4.2 Gaussian representation and the average over H

We now write the superdeterminant in (4.2) as a Gaussian integral, as in Lemma 3.9, and compute the average over H. For each site $i \in [N]$, we take $2k$ complex coordinates $\zeta_i = (\zeta_{i,1}, \dots, \zeta_{i,2k})$ and $4k$ odd coordinates $\chi_{i,a}, \bar{\chi}_{i,a}$, $a \in [2k]$, and we let $\psi_i = (\zeta_i, \chi_i)$ denote the supervector at the site i. As in Definition 3.7, the integral over all these coordinates is denoted $\int \cdot \,\mathrm{d}\psi$. For a $(2k|2k)$ supermatrix M, we set $\psi_i^* M \psi_j = (\zeta_i, \chi_i)^* M (\zeta_j, \chi_j)$, in the notation preceding Proposition 3.12, and we let $\psi_i\psi_j^*$ denote the even supermatrix with entries $(\psi_i)_a(\psi_j^*)_b$. One checks on the four blocks that

$$\operatorname{STr}(\psi_i\psi_j^* M) = \psi_j^* M \psi_i \tag{4.4}$$

for every even supermatrix M; in the FF block, we used $-\chi_{i,a}\bar{\chi}_{j,b} = \bar{\chi}_{j,b}\chi_{i,a}$.

Lemma 4.4. *Let Γ be a source supermatrix and Y local sources with $\|Y^{\mathrm{B}}\|_{\mathrm{op}} < \eta$. Then*

$$T_N^{(k)}(\Gamma; Y) = \int \exp\Big(\mathrm{i}\sum_{i=1}^{N} \psi_i^*\Lambda(E + \mathrm{i}\eta_0\Gamma + Y_i)\psi_i - \frac{1}{2}\sum_{i,j=1}^{N} J_{ij}\operatorname{STr}\left(\psi_i\psi_i^*\Lambda\,\psi_j\psi_j^*\Lambda\right)\Big)\,\mathrm{d}\psi, \tag{4.5}$$

where the integrand is integrable.

Proof. Fix a Hermitian matrix H, and set

$$M = -\mathrm{i}(\mathrm{Id}_N \otimes \Lambda)(E + \mathrm{i}\eta_0\Gamma + Y - H),$$

which has vanishing odd blocks. The proof of Lemma 4.3 shows that

$$\operatorname{Re} M^{\mathrm{BB}} = \operatorname{Im}\big((\mathrm{Id}_N \otimes \Lambda_{\mathrm{B}})(E + \mathrm{i}\eta_0\Gamma + Y - H)^{\mathrm{BB}}\big) \geq (\eta - \|Y^{\mathrm{B}}\|_{\mathrm{op}})\mathrm{Id}.$$

Note that $\operatorname{SDet}(-\mathrm{i}\,\mathrm{Id}_N \otimes \Lambda) = 1$, since its BB and FF blocks coincide. By Proposition 3.12, we have

$$\operatorname{SDet}\big(E + \mathrm{i}\eta_0\Gamma + Y - H\big)^{-1} = \int \exp\Big(\mathrm{i}\sum_i \psi_i^*\Lambda(E + \mathrm{i}\eta_0\Gamma + Y_i)\psi_i - \mathrm{i}\sum_{i,j} H_{ij}\,\psi_i^*\Lambda\psi_j\Big)\,\mathrm{d}\psi.$$

The body of this integrand has modulus

$$\exp\Big(-\sum_i \zeta_i^* \operatorname{Im}\big(\Lambda_{\mathrm{B}}(\mathrm{i}\eta\Lambda_{\mathrm{B}} + Y_i^{\mathrm{B}})\big)\zeta_i\Big) \leq e^{-(\eta - \|Y^{\mathrm{B}}\|_{\mathrm{op}})|\zeta|^2},$$

which is independent of H, and the other coefficients are this function times polynomials in ζ, $\bar{\zeta}$, and the entries of H. Using Fubini's theorem, we may take the expectation inside the integral.

We now compute the expectation. Let $c = (c_{ij})$ be a complex matrix. The variable $\sum_{i,j} H_{ij}c_{ij}$ is a complex linear combination of centered jointly Gaussian variables. By (1.1), we have

$$\mathbb{E}\Big(\sum_{i,j} H_{ij}c_{ij}\Big)^2 = \sum_{i,j} J_{ij}c_{ij}c_{ji}, \qquad \mathbb{E}\exp\Big(-\mathrm{i}\sum_{i,j} H_{ij}c_{ij}\Big) = \exp\Big(-\frac{1}{2}\sum_{i,j} J_{ij}c_{ij}c_{ji}\Big).$$

Each side of the second identity is entire in c. Using Lemma 3.4(c), we conclude that the identity holds for the even elements $c_{ij} = \psi_i^*\Lambda\psi_j$. Finally, since $\psi_i^*\Lambda\psi_j$ is even, we have

$$\operatorname{STr}\big(\psi_i\psi_i^*\Lambda\psi_j\psi_j^*\Lambda\big) = (\psi_i^*\Lambda\psi_j)\operatorname{STr}\big(\psi_i\psi_j^*\Lambda\big) = (\psi_i^*\Lambda\psi_j)(\psi_j^*\Lambda\psi_i),$$

where we used (4.4) in the last step. □

The quadratic term in (4.5) contains the spectral parameters and is a sum of terms attached to single sites. The quartic term comes from the randomness of H. It couples every pair of sites with the weight J_{ij}, and by (1.2) the weight is of order W^{-1} for pairs at distance up to order W. The rest of the derivation removes this coupling.

4.3 Decoupling and the nearest-neighbor chain

The Hubbard–Stratonovich transformation decouples a quadratic form in the exponent by a Gaussian integral. In one variable, it is the identity $e^{-a^2/2} = (2\pi)^{-1/2}\int_{\mathbb{R}} e^{-r^2/2 - \mathrm{i}ar}\,\mathrm{d}r$.

Given a positive definite matrix J and a vector a, one uses instead a centered Gaussian vector g with covariance J,

$$e^{-\langle a,Ja\rangle/2} = \mathbb{E}e^{-\mathrm{i}\langle a,g\rangle}, \qquad \text{where } g \text{ has the density } \frac{e^{-\langle r,J^{-1}r\rangle/2}}{(2\pi)^{N/2}(\det J)^{1/2}}.$$

We apply this with the supermatrix $\psi_i\psi_i^*\Lambda$ in place of a_i, and with one supermatrix-valued variable R_i for each site in place of g_i. Each R_i is a $(2k|2k)$ supermatrix, integrated over a cycle. We use this word as in contour integration. A *cycle* is a set of $(2k|2k)$ supermatrices whose odd entries are independent odd coordinates and whose pairs of BB and FF blocks range over a real submanifold of $\mathbb{C}^{2k\times 2k}\times\mathbb{C}^{2k\times 2k}$ of real dimension $8k^2$, which is the complex dimension of this space. The integrands below are holomorphic in the BB and FF entries, and in Section 4.4 we move the cycle in complex space, as one moves a contour of integration in one complex variable. We begin with the flat cycle.

Definition 4.5. We let $\mathcal{S}_\flat$ denote the set of $(2k|2k)$ supermatrices whose BB block is Hermitian, whose FF block is $\mathrm{i}M$ with M Hermitian, and whose $8k^2$ odd entries are independent odd coordinates. We identify $\mathcal{S}_\flat$ with the superdomain $\mathbb{R}^{8k^2|8k^2}$, through the real coordinates of the BB block and of M, orthonormal for the Hilbert–Schmidt inner product, and the odd entries, ordered so that R^{BF}_{ab} is immediately followed by R^{FB}_{ba} for all $a,b\in[2k]$. We let $\mathrm{d}R$ denote the Berezin integral of Definition 3.5 on this superdomain, multiplied by the constant $c_k\in\mathbb{C}\setminus\{0\}$ for which

$$\int_{\mathcal{S}_\flat} e^{-\operatorname{STr}(R^2)/2}\,\mathrm{d}R = 1. \tag{4.6}$$

On $\mathcal{S}_\flat^N$, we use the product of N copies of $\mathrm{d}R$, with independent odd coordinates.

Note that on $\mathcal{S}_\flat$,

$$\operatorname{STr} R^2 = \|R^{\mathrm{BB}}\|^2 + \|M\|^2 + 2\operatorname{Tr}\left(R^{\mathrm{BF}}R^{\mathrm{FB}}\right),$$

where we used $\operatorname{Tr}(R^{\mathrm{FB}}R^{\mathrm{BF}}) = -\operatorname{Tr}(R^{\mathrm{BF}}R^{\mathrm{FB}})$, since the entries of the odd blocks anticommute. We now compute c_k. The even coordinates contribute the Gaussian integral $(2\pi)^{4k^2}$ over $\mathbb{R}^{8k^2}$. The products $R^{\mathrm{BF}}_{ab}R^{\mathrm{FB}}_{ba}$ are even and commute, and each pair of their factors is consecutive in our order of the odd entries. Applying Lemma 3.6(f) pair by pair, we find

$$\int e^{-\operatorname{Tr}(R^{\mathrm{BF}}R^{\mathrm{FB}})}\,\mathrm{d}\xi = \prod_{a,b=1}^{2k}\int\left(1-R^{\mathrm{BF}}_{ab}R^{\mathrm{FB}}_{ba}\right)\mathrm{d}R^{\mathrm{BF}}_{ab}\,\mathrm{d}R^{\mathrm{FB}}_{ba} = (-1)^{4k^2} = 1,$$

where ξ denotes the $8k^2$ odd entries of R in this order. Therefore $c_k = (2\pi)^{-4k^2}$. In another order of the odd entries, this Berezin integral may be -1; for instance, reading the two odd blocks row by row gives $(-1)^k$.

The flat cycle has the three features of the GUE computation listed after the proof of Proposition 3.13: the Gaussian measure in (4.7) below has total integral one; the FF block is rotated by i, since STr carries a minus sign on the FF block; and the BB block is Hermitian, which suffices only when one half plane appears (see Section 4.4). We prove in Lemma B.2 that for every positive definite real symmetric $N \times N$ matrix $\mathbf{M}$ and all even supermatrices $B_1, \dots, B_N$ with entries in auxiliary generators,

$$\exp\Big(-\frac{1}{2}\sum_{i,j}(\mathbf{M}^{-1})_{ij}\,\mathrm{STr}(B_iB_j)\Big) = \int_{\mathcal{S}_\flat^N} \exp\Big(-\frac{1}{2}\sum_{i,j}\mathbf{M}_{ij}\,\mathrm{STr}(R_iR_j) - \mathrm{i}\sum_i \mathrm{STr}(R_iB_i)\Big)\,\mathrm{d}R. \tag{4.7}$$

The right side carries the constant one for all $\mathbf{M}$, and this is the basic supersymmetric cancellation. We spell it out for $B = 0$, to which the general case reduces by completing the square. Let $x_i \in \mathbb{R}^{8k^2}$ denote the even coordinates of R_i. Since $\mathbf{M}$ is symmetric, we find as in the computation of $\mathrm{STr}\,R^2$ above that

$$\frac{1}{2}\sum_{i,j=1}^N \mathbf{M}_{ij}\,\mathrm{STr}(R_iR_j) = \frac{1}{2}\sum_{i,j=1}^N \mathbf{M}_{ij}\langle x_i, x_j\rangle + \sum_{a,b=1}^{2k}\sum_{i,j=1}^N (R_i^{\mathrm{BF}})_{ab}\,\mathbf{M}_{ij}\,(R_j^{\mathrm{FB}})_{ba}.$$

For the first term, the Gaussian integral over the even coordinates is

$$\begin{aligned}\int_{\mathbb{R}^{8k^2N}} \exp\Big(-\frac{1}{2}\sum_{i,j=1}^N \mathbf{M}_{ij}\langle x_i, x_j\rangle\Big)\,\mathrm{d}x &= (2\pi)^{4k^2N}\det(\mathbf{M}\otimes \mathrm{Id}_{8k^2})^{-1/2}\\ &= (2\pi)^{4k^2N}\det(\mathbf{M})^{-4k^2}.\end{aligned}$$

For each pair (a,b), the second term is a fermionic Gaussian in the $2N$ odd entries $(R_i^{\mathrm{BF}})_{ab}$ and $(R_i^{\mathrm{FB}})_{ba}$, and since the odd entries of each R_i come in consecutive pairs, the Berezin integral factors over (a,b) without signs. By Proposition 3.8(b), applied with $\chi_i = (R_i^{\mathrm{BF}})_{ab}$, $\bar\chi_i = (R_i^{\mathrm{FB}})_{ba}$, and the matrix $-\mathbf{M}$, each pair contributes the factor $\det(-\mathbf{M}) = (-1)^N \det\mathbf{M}$. Over the $4k^2$ pairs, we obtain

$$\int_{\mathcal{S}_\flat^N} e^{-\frac{1}{2}\sum_{i,j}\mathbf{M}_{ij}\,\mathrm{STr}(R_iR_j)}\,\mathrm{d}R = c_k^N\,(2\pi)^{4k^2N}\det(\mathbf{M})^{-4k^2}\big((-1)^N\det\mathbf{M}\big)^{4k^2} = 1.$$

The even coordinates contribute $\det(\mathbf{M})^{-4k^2}$ and the odd coordinates $\det(\mathbf{M})^{4k^2}$, and the two factors cancel because $\mathcal{S}_\flat$ has as many even as odd dimensions, $8k^2$ of each.

To decouple (4.5), we take $\mathbf{M} = J^{-1}$ and $B_i = \psi_i\psi_i^*\Lambda$. This is where the choice of the variance profile matters. Recall that $J^{-1} = \mathrm{Id} - W^2\Delta$, and note that

$$\sum_{i,j=1}^N (-\Delta)_{ij}f(i)f(j) = \sum_{i=1}^{N-1}\big(f(i+1)-f(i)\big)^2.$$

Using this identity in each entry and $\mathrm{STr}(AB) = \mathrm{STr}(BA)$ for even supermatrices, we obtain

$$\sum_{i,j=1}^{N} (J^{-1})_{ij} \,\mathrm{STr}(R_i R_j) = \sum_{i=1}^{N} \mathrm{STr}\, R_i^2 + W^2 \sum_{i=1}^{N-1} \mathrm{STr}(R_{i+1} - R_i)^2. \tag{4.8}$$

Thus the precision of the Hubbard–Stratonovich field is a nearest-neighbor operator, although J itself couples all pairs of sites. This is the reason for the choice of J in Definition 1.1. The same profile is used in [29], where it is observed that the dual supersymmetric model has nearest-neighbor coupling, and in three dimensions in [27].

We now insert (4.7) into (4.5) and, for the moment formally, integrate out ψ. By (4.4), we have $\mathrm{STr}(R_i \psi_i \psi_i^* \Lambda) = \psi_i^* \Lambda R_i \psi_i$. The exponent becomes $\mathrm{i} \sum_i \psi_i^* \Lambda (E + \mathrm{i}\eta_0 \Gamma + Y_i - R_i) \psi_i$, and the ψ-integral factors over the sites. Applying Proposition 3.12 at each site, with $Y = 0$, we would obtain

$$T_N^{(k)}(\Gamma) \stackrel{?}{=} \int_{\mathcal{S}_\flat^N} \exp\Big(- \sum_{i=1}^{N} \Big[\frac{1}{2} \mathrm{STr}\, R_i^2 + \mathrm{STr} \log \big(E + \mathrm{i}\eta_0 \Gamma - R_i \big) \Big] - \frac{W^2}{2} \sum_{i=1}^{N-1} \mathrm{STr}(R_{i+1} - R_i)^2 \Big) \,\mathrm{d}R, \tag{4.9}$$

where we wrote $\mathrm{SDet}(A)^{-1} = e^{-\mathrm{STr} \log A}$, which follows from Lemma 3.11(d) when the logarithm is defined. The source appears only in the onsite terms. We translate $R_i = S_i + \mathrm{i}\eta_0 \Gamma$, which leaves the links unchanged, and we expand

$$\frac{1}{2} \mathrm{STr}(S_i + \mathrm{i}\eta_0 \Gamma)^2 = \frac{1}{2} \mathrm{STr}\, S_i^2 + \mathrm{i}\eta_0 \,\mathrm{STr}(S_i \Gamma) - \frac{1}{2} \eta_0^2 \,\mathrm{STr}\, \Gamma^2.$$

The right side of (4.9) then becomes

$$e^{N\eta_0^2 \,\mathrm{STr}(\Gamma^2)/2} \int \prod_{i=1}^{N} e^{-\mathrm{STr}(S_i^2)/2 - \mathrm{i}\eta_0 \,\mathrm{STr}(S_i \Gamma)} \,\mathrm{SDet}(E - S_i)^{-1} \prod_{i=1}^{N-1} e^{-W^2 \,\mathrm{STr}(S_{i+1} - S_i)^2/2} \,\mathrm{d}S. \tag{4.10}$$

The integrand in (4.10) is a product of N factors attached to single sites and $N - 1$ Gaussian links with variance of order W^{-2} in $S_{i+1} - S_i$, so that the integral is a matrix element of the $(N-1)$st power of an integral operator in one supermatrix variable, which we define in Section 4.5 (see Figure 4.1). Recall from the discussion after (2.2) that the resemblance of such an integral to a Feynman–Kac formula is only formal, since the weights are complex and contain odd variables.

For other variance profiles, the first two steps are unchanged, since Lemma 4.4 uses only the variances (1.1). However, the decoupling by one supermatrix per site needs a Gaussian field with covariance J, and the precision J^{-1} of this field appears in the action. For the sharp profile $J_{ij} = (2W+1)^{-1} \mathbf{1}_{|i-j| \le W}$, such a field does not exist once N is large compared with W, since J is then not positive semidefinite. Indeed, its symbol,

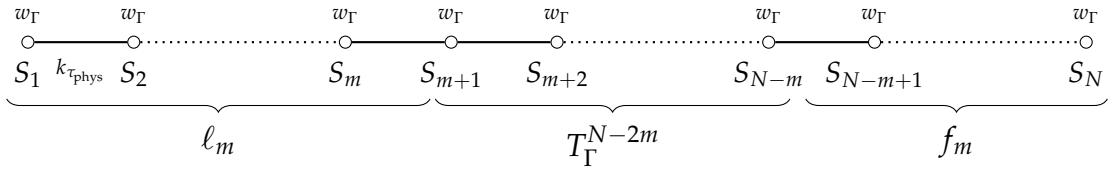


Figure 4.1. The open chain of Proposition 4.11. Each site carries the weight w_Γ of (4.15), and each pair of neighboring sites the link kernel $k_{\tau_{\text{phys}}}$ of Definition 4.9. In the pairing (4.19), the partial chain f_m collects the last m sites, the power T_Γ^{N-2m} the middle $N-2m$ sites, and ℓ_m the first m sites together with the link from S_m to S_{m+1}.

the normalized Dirichlet kernel, takes negative values, and for $W = 8$ and $N = 200$ the smallest eigenvalue of J is about -0.22. A positive definite profile whose inverse is not sparse leads to a decoupled integrand that contains $\mathrm{STr}(R_i R_j)$ for all pairs of sites and does not factor into nearest-neighbor terms. We do not know how to write the chain for such profiles with an integral operator in one supermatrix variable, and our proof uses the nearest-neighbor structure (4.8) in an essential way.

4.4 The noncompact sector and the Schäfer–Wegner cycle

The right side of (4.9) is not defined as written. We now explain why, and how the derivation is repaired. The identity (4.7) holds for each fixed ψ. To integrate out ψ afterward, we must exchange the ψ-integral with the integral over R, and for this the integrand must be integrable on the product. The odd variables appear polynomially, and the FF blocks through convergent Gaussians. The difficulty is in the BB blocks. For $R_i \in \mathcal{S}_\flat$, the commuting part of the exponent at the site i has the real part

$$\mathrm{Re}\left[\mathrm{i}\zeta_i^* \Lambda_{\mathrm{B}}(E + \mathrm{i}\eta\Lambda_{\mathrm{B}} - R_i^{\mathrm{BB}})\zeta_i\right] = -\eta|\zeta_i|^2 + \zeta_i^* \,\mathrm{Im}\left(\Lambda_{\mathrm{B}} R_i^{\mathrm{BB}}\right)\zeta_i.$$

Since R_i^{BB} is Hermitian, $\mathrm{Im}(\Lambda_{\mathrm{B}} R_i^{\mathrm{BB}})$ is a Hermitian matrix with trace zero, whose eigenvalues are of the order of the off-diagonal blocks of R_i^{BB}. The ζ_i-integral therefore diverges whenever these blocks are large compared with η. Equivalently, the inverse superdeterminant in (4.9) has poles on $\mathcal{S}_\flat$. For instance, for $k = 1$ and the Hermitian matrix $R^{\mathrm{BB}} = \left(\begin{smallmatrix} E & r \\ \bar{r} & E \end{smallmatrix}\right)$, we have

$$\det\left(E + \mathrm{i}\eta\Lambda_{\mathrm{B}} - R^{\mathrm{BB}}\right) = \eta^2 - |r|^2,$$

which vanishes for $|r| = \eta$.

The source of the difficulty is that the quadratic form $\zeta^* \Lambda_{\mathrm{B}} \zeta$ is indefinite, of signature (k, k), so that the symmetry group of the commuting variables is the noncompact group $U(k, k)$. This hyperbolic symmetry was first observed by Wegner [134] (see also [60, p. 2]), and the cure is due to Schäfer and Wegner [111]. One deforms the BB blocks of the Hubbard–Stratonovich variables into the complex domain, in such a way that

$\mathrm{Im}(\Lambda_{\mathrm{B}}(E + \mathrm{i}\eta\Lambda_{\mathrm{B}} - R_i^{\mathrm{BB}}))$ becomes positive definite. We use the following form of their cycle. We let $\mathrm{Herm}(2k)$ denote the space of Hermitian $2k \times 2k$ matrices, and we let $\mathrm{Herm}_{\mathrm{d}}(2k)$ and $\mathrm{Herm}_{\mathrm{o}}(2k)$ denote its subspaces of block-diagonal and block off-diagonal matrices, for the decomposition into $k \times k$ blocks given by the retarded and advanced flavors.

From here on, in this section, in Sections 15 and 19, and in Appendix B, the letter X denotes a matrix in $\mathrm{Herm}_{\mathrm{o}}(2k)$, which will be a coordinate of the cycle. The parameter X of Definition 1.2 appears there only as $X_N(E)$.

For $X \in \mathrm{Herm}_{\mathrm{o}}(2k)$, we set

$$f_0(X) = \big(s_E^2\mathrm{Id} + X^2\big)^{1/2}, \qquad \mathrm{g}_X = \big(s_E^{-1}(f_0(X) + X)\big)^{1/2}. \tag{4.11}$$

Definition 4.6. We let the *Schäfer–Wegner cycle* $\mathcal{S}$ be the set of $(2k|2k)$ supermatrices S with

$$S^{\mathrm{BB}} = P - \mathrm{i}\Lambda_{\mathrm{B}}\big(f_0(X) + X\big), \qquad S^{\mathrm{FF}} = \kappa_E\mathrm{Id} + \mathrm{i}M, \tag{4.12}$$

where $P \in \mathrm{Herm}_{\mathrm{d}}(2k)$, $X \in \mathrm{Herm}_{\mathrm{o}}(2k)$, and $M \in \mathrm{Herm}(2k)$, and whose odd entries are independent odd coordinates. We call (P, X, M) the coordinates of $S \in \mathcal{S}$. For $R \in \mathcal{S}_\flat$ with BB block $P - \mathrm{i}\Lambda_{\mathrm{B}}X$ and FF block $\mathrm{i}M$, we let $S(R)$ denote the point of $\mathcal{S}$ with coordinates (P, X, M) and the same odd entries as R. For a superfunction F of S whose coefficients are defined on $\mathcal{S}$, we call F *integrable on* $\mathcal{S}$ if $R \mapsto F(S(R))$ is integrable in the sense of Definition 3.5, and we define

$$\int_{\mathcal{S}} F(S)\,\mathrm{d}S = \int_{\mathcal{S}_\flat} F(S(R))\,\mathrm{d}R.$$

We let $\mathcal{S}^N$ denote the product of N copies of $\mathcal{S}$, with independent odd coordinates, and the product integral.

The map $X \mapsto -\mathrm{i}\Lambda_{\mathrm{B}}X$ is a linear isometry of $\mathrm{Herm}_{\mathrm{o}}(2k)$ onto itself. Every Hermitian matrix therefore has the form $P - \mathrm{i}\Lambda_{\mathrm{B}}X$ for unique P and X, and the flat cycle has BB block $P - \mathrm{i}\Lambda_{\mathrm{B}}X$. The cycle $\mathcal{S}$ is thus a graph over $\mathcal{S}_\flat$: its BB block differs from that of $\mathcal{S}_\flat$ by the block-diagonal matrix $-\mathrm{i}\Lambda_{\mathrm{B}}f_0(X)$, and its FF block by the constant $\kappa_E\mathrm{Id}$. Figure 4.2 shows this graph for $k = 1$.

The definition of $\int_{\mathcal{S}} F\,\mathrm{d}S$ contains no Jacobian factor, because it is a contour integral of a holomorphic integrand. To see this, we write the BB block of a complex supermatrix as $P^{\mathbb{C}} - \mathrm{i}\Lambda_{\mathrm{B}}X^{\mathbb{C}}$, with $P^{\mathbb{C}}$ complex block diagonal and $X^{\mathbb{C}}$ complex block off-diagonal, and its FF block as $\mathrm{i}M^{\mathbb{C}}$. We use the entries of $P^{\mathbb{C}}$, $X^{\mathbb{C}}$, and $M^{\mathbb{C}}$ in the complexifications of orthonormal bases of the corresponding real spaces as complex coordinates $z \in \mathbb{C}^{8k^2}$. These coordinates are real exactly on $\mathcal{S}_\flat$, where they are the coordinates of Definition 4.5. The form $\mathrm{d}R$ then extends to the *holomorphic Berezin form* $c_k\,\mathrm{d}z_1 \wedge \cdots \wedge \mathrm{d}z_{8k^2}\,\mathrm{d}\xi$, the product of a holomorphic volume form in the even entries and the Berezin integral in

the odd entries. As for a contour integral $\int_\gamma f(z)\,\mathrm{d}z = \int f(\gamma(t))\gamma'(t)\,\mathrm{d}t$, integrating this form over a cycle parametrized by $x \mapsto z(x)$ means integrating $F(z(x))\det z'(x)$ against $c_k\,\mathrm{d}x\,\mathrm{d}\xi$, where $z'(x)$ denotes the complex $8k^2 \times 8k^2$ matrix of partial derivatives of z in x. By Lemma B.1, the result does not change when the cycle is deformed, provided the integrand decays. The cycle $\mathcal{S}$ is parametrized by

$$(P, X, M) \mapsto (P^{\mathbb{C}}, X^{\mathbb{C}}, M^{\mathbb{C}}) = \big(P - \mathrm{i}\Lambda_{\mathrm{B}} f_0(X),\ X,\ M - \mathrm{i}\kappa_E \mathrm{Id}\big),$$

whose Jacobian matrix is triangular with unit diagonal. Hence $\det z' = 1$, and $\int_{\mathcal{S}} F\,\mathrm{d}S$ is the integral of the holomorphic Berezin form over $\mathcal{S}$.

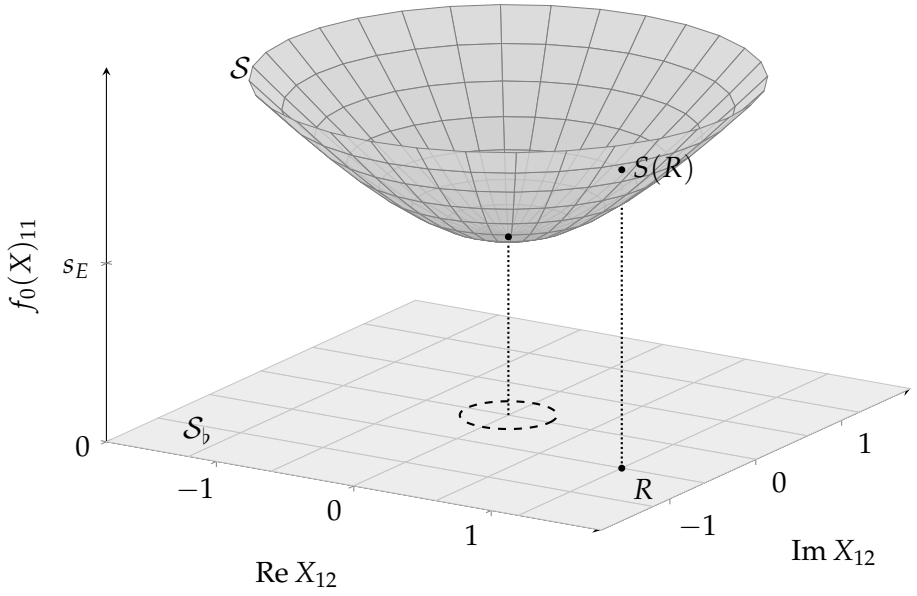


Figure 4.2. The Schäfer–Wegner cycle for $k = 1$, drawn for $E = 0$ and $P = 0$, so that $s_E = 1$ and $P = E\mathrm{Id} = \kappa_E \mathrm{Id}$. Here X has the single entry $X_{12} = \bar{X}_{21}$. The plane of X_{12} at height zero carries $\mathcal{S}_\flat$, and $\mathcal{S}$ is the graph of $f_0(X)_{11} = -\operatorname{Im} S_{11}^{\mathrm{BB}}$ over it, the upper sheet of the hyperboloid $f_0(X)_{11}^2 - |X_{12}|^2 = s_E^2$. The two dotted segments join points R of $\mathcal{S}_\flat$ to the points $S(R)$ above them, one at $X = 0$, where $S(R)$ is the vertex at height s_E, and one at the labeled point R. The dashed circle $|X_{12}| = \eta$, drawn with $\eta = 0.3$, carries the poles on $\mathcal{S}_\flat$ of the example in Section 4.4, with $|r| = |X_{12}|$, and $\mathcal{S}$ stays away from it. The points of $\mathcal{S}$ shown have the BB blocks $\kappa_E - \mathrm{i}s_E Q_X$ of Lemma 4.7(d), the BB blocks of critical points of the weight of the chain, and the vertex gives $\kappa_E - \mathrm{i}s_E\Lambda_{\mathrm{B}}$.

Since P ranges over all of $\mathrm{Herm}_{\mathrm{d}}(2k)$, the cycle $\mathcal{S}$ is also the set of supermatrices with BB block $\kappa_E\mathrm{Id} + P - \mathrm{i}\Lambda_{\mathrm{B}}(f_0(X) + X)$, which is the form used in Part IV. We will need the following properties of $\mathcal{S}$.

Lemma 4.7. *Let $X \in \mathrm{Herm}_{\mathrm{o}}(2k)$ and $S \in \mathcal{S}$ with coordinates (P, X, M).*

(a) The matrix $f_0(X)$ is block diagonal and commutes with X, the matrix $\mathrm{g}_X^2 = s_E^{-1}(f_0(X)+X)$ is positive definite with $\det \mathrm{g}_X = 1$ and $\mathrm{g}_X \Lambda_{\mathrm{B}} \mathrm{g}_X = \Lambda_{\mathrm{B}}$, and $\operatorname{Tr} f_0(X) \geq \|X\|$.

(b) We have $\operatorname{Im}\big(\Lambda_{\mathrm{B}}(E - S^{\mathrm{BB}})\big) = s_E \mathrm{g}_X^2$. In particular $E - S^{\mathrm{BB}}$ is invertible, with

$$|\det(E - S^{\mathrm{BB}})|^{-1} \leq s_E^{-2k}, \qquad \|(E - S^{\mathrm{BB}})^{-1}\|_{\mathrm{op}} \leq s_E^{-2}\big(s_E + 2\|X\|_{\mathrm{op}}\big).$$

(c) The body of $\operatorname{STr} S^2$ has real part $\|P\|^2 - 2ks_E^2 + \|M\|^2 - 2k\kappa_E^2$, and for $\theta \in [0,1]$,

$$\operatorname{Re}\Big[-\frac{1}{2} \operatorname{Tr}\big(S^{\mathrm{BB}} + \mathrm{i}\theta\eta\Lambda_{\mathrm{B}}\big)^2 \Big] = -\frac{1}{2}\|P\|^2 - \theta\eta \operatorname{Tr} f_0(X) + k\big(s_E^2 + \theta^2\eta^2\big). \tag{4.13}$$

(d) For $Q_X = \mathrm{g}_X^{-1}\Lambda_{\mathrm{B}}\mathrm{g}_X$, we have $Q_X^2 = \mathrm{Id}$ and $-\mathrm{i}\Lambda_{\mathrm{B}}(f_0(X)+X) = -\mathrm{i}s_E Q_X$. In particular, the point of $\mathcal{S}$ with $P = \kappa_E \mathrm{Id}$ has BB block $\kappa_E \mathrm{Id} - \mathrm{i}s_E Q_X$.

Proof. We begin with (a). The matrices X^2 and $f_0(X)$ are block diagonal, since X is block off-diagonal. The identity $\Lambda_{\mathrm{B}} X \Lambda_{\mathrm{B}} = -X$ shows that the eigenvalues of X come in pairs $\pm x$, and $f_0(X) + X$ has the eigenvalues $(s_E^2 + x^2)^{1/2} + x > 0$. Using

$$\big((s_E^2 + x^2)^{1/2} + x\big)\big((s_E^2 + x^2)^{1/2} - x\big) = s_E^2,$$

we find that $\det \mathrm{g}_X^2 = 1$. Moreover

$$\Lambda_{\mathrm{B}} \mathrm{g}_X^2 \Lambda_{\mathrm{B}} = s_E^{-1}\big(f_0(X) - X\big) = \mathrm{g}_X^{-2},$$

and taking positive square roots, we obtain $\Lambda_{\mathrm{B}} \mathrm{g}_X \Lambda_{\mathrm{B}} = \mathrm{g}_X^{-1}$, which is equivalent to $\mathrm{g}_X \Lambda_{\mathrm{B}} \mathrm{g}_X = \Lambda_{\mathrm{B}}$. Finally,

$$\operatorname{Tr} f_0(X) = \sum_x (s_E^2 + x^2)^{1/2} \geq \Big(\sum_x x^2 \Big)^{1/2} = \|X\|.$$

For (b), we write

$$\Lambda_{\mathrm{B}}(E - S^{\mathrm{BB}}) = \Lambda_{\mathrm{B}}(E - P) + \mathrm{i}\big(f_0(X) + X\big) = \Lambda_{\mathrm{B}}(E - P) + \mathrm{i}s_E \mathrm{g}_X^2.$$

Since P commutes with Λ_{B}, the matrix $\Lambda_{\mathrm{B}}(E - P)$ is Hermitian, and the formula for the imaginary part follows. The bound on the determinant follows from the inequality $|\det A| \geq \det \operatorname{Im} A$ of the proof of Lemma 4.3 and $\det(s_E \mathrm{g}_X^2) = s_E^{2k}$.

To bound the inverse, let A be a matrix with $\operatorname{Im} A > 0$. For all unit vectors v, we have

$$\|Av\| \geq |v^* A v| \geq v^* (\operatorname{Im} A) v,$$

which implies that

$$\|A^{-1}\|_{\mathrm{op}} \leq \|(\operatorname{Im} A)^{-1}\|_{\mathrm{op}}.$$

We apply this with $A = \Lambda_{\mathrm{B}}(E - S^{\mathrm{BB}})$, and we obtain

$$\big\|(E - S^{\mathrm{BB}})^{-1}\big\|_{\mathrm{op}} \leq s_E^{-1}\|\mathrm{g}_X^{-2}\|_{\mathrm{op}} = s_E^{-2}\|f_0(X) - X\|_{\mathrm{op}} \leq s_E^{-2}\big(s_E + 2\|X\|_{\mathrm{op}}\big),$$

where we used $\mathrm{g}_X^{-2} = s_E^{-1}(f_0(X) - X)$ and $\|f_0(X)\|_{\mathrm{op}} \leq s_E + \|X\|_{\mathrm{op}}$.

For (c), we set $Z = f_0(X) + X - \theta\eta$, which is Hermitian, so that

$$S^{\mathrm{BB}} + \mathrm{i}\theta\eta\Lambda_{\mathrm{B}} = P - \mathrm{i}\Lambda_{\mathrm{B}}Z.$$

The trace $\mathrm{Tr}(P\Lambda_{\mathrm{B}}Z) = \mathrm{Tr}((\Lambda_{\mathrm{B}}P)Z)$ is real, since P and Λ_{B} commute. The cross terms between the block-diagonal and the block off-diagonal components of Z have trace zero, and

$$\Lambda_{\mathrm{B}}Z\Lambda_{\mathrm{B}} = f_0(X) - \theta\eta - X.$$

Using $\|f_0(X)\|^2 = 2ks_E^2 + \|X\|^2$, we find

$$\begin{aligned}\mathrm{Re}\,\mathrm{Tr}\,\big(P - \mathrm{i}\Lambda_{\mathrm{B}}Z\big)^2 &= \|P\|^2 + \|X\|^2 - \|f_0(X) - \theta\eta\|^2 \\ &= \|P\|^2 - 2ks_E^2 + 2\theta\eta\,\mathrm{Tr}\,f_0(X) - 2k\theta^2\eta^2.\end{aligned}$$

This is (4.13), and the case $\theta = 0$ gives the BB part of the first claim. The FF block contributes

$$-\,\mathrm{Re}\,\mathrm{Tr}(\kappa_E + \mathrm{i}M)^2 = \|M\|^2 - 2k\kappa_E^2,$$

and the odd entries contribute only nilpotent terms.

For (d), note that part (a) implies $\Lambda_{\mathrm{B}}\mathrm{g}_X = \mathrm{g}_X^{-1}\Lambda_{\mathrm{B}}$, and therefore

$$\Lambda_{\mathrm{B}}\big(f_0(X) + X\big) = s_E\Lambda_{\mathrm{B}}\mathrm{g}_X^2 = s_E\mathrm{g}_X^{-1}\Lambda_{\mathrm{B}}\mathrm{g}_X. \qquad \square$$

The cycle is designed for Lemma 4.7(b). For $R_i \in \mathcal{S} + \mathrm{i}\eta_0\Gamma$, the BB block of $\Lambda(E + \mathrm{i}\eta_0\Gamma - R_i)$ has imaginary part $s_E\mathrm{g}_{X_i}^2 > 0$. This makes the ψ_i-integral converge and shows that the inverse superdeterminant in (4.10) is a smooth function on $\mathcal{S}$. Next, Lemma 4.7(d) shows that the BB blocks of $\mathcal{S}$ include the matrices $\kappa_E - \mathrm{i}s_E g\Lambda_{\mathrm{B}}g^{-1}$ with $g > 0$ and $g\Lambda_{\mathrm{B}}g = \Lambda_{\mathrm{B}}$, and the FF blocks of $\mathcal{S}$ include the matrices $\kappa_E - \mathrm{i}s_E Q_{\mathrm{F}}$ with Q_{F} Hermitian and $Q_{\mathrm{F}}^2 = \mathrm{Id}$. These are the blocks of the critical points of the weight of the chain near which the chain concentrates, which we describe in Section 4.6. This is the reason for the choice $s = s_E$ in f_0.

The integrability on the deformed cycle is delicate, and (4.13) shows why. Consider first the cycle $\mathcal{S}$ itself, which is the case $\theta = 0$. There the modulus of the BB part of the Gaussian factor $e^{-\mathrm{STr}(R^2)/2}$ is constant in X, so that this factor does not decay in the noncompact directions. The decay comes from the translation by $\mathrm{i}\eta_0\Gamma$, which is the case $\theta = 1$ and produces the factor $e^{-\eta\,\mathrm{Tr}\,f_0(X)}$. In the variables $S_i = R_i - \mathrm{i}\eta_0\Gamma$, this factor is the BB part of $|e^{-\mathrm{i}\eta_0\,\mathrm{STr}(S\Gamma)}|$, the damping factor of the weight that each site of the

chain carries (Definition 4.9). For this reason, we deform the flat cycle directly onto the translated cycle $\mathcal{S} + \mathrm{i}\eta_0\Gamma$, through the cycles $\mathcal{S}_{\Gamma,t}$, $t \in [0,1]$, with BB blocks

$$P - \mathrm{i}\Lambda_{\mathrm{B}}\big((1-t)(f_0(X)-\eta)+X\big), \qquad P \in \mathrm{Herm_d}(2k),\ X \in \mathrm{Herm_o}(2k),$$

FF blocks $\mathrm{i}M + (1-t)(\kappa_E \mathrm{Id} + \mathrm{i}\eta_0\Gamma_{\mathrm{F}})$ with $M \in \mathrm{Herm}(2k)$, and independent odd entries. Thus $\mathcal{S}_{\Gamma,1} = \mathcal{S}_\flat$ and $\mathcal{S}_{\Gamma,0} = \mathcal{S} + \mathrm{i}\eta_0\Gamma$. The integral over $\mathcal{S}_{\Gamma,t}$ is defined as in Definition 4.6, and $\mathcal{S}_{\Gamma,t}^N$ denotes the product cycle. We will need the following identity.

Proposition 4.8. *Let Γ be a source supermatrix, and fix $\zeta_1, \dots, \zeta_N \in \mathbb{C}^{2k}$. Let $F_\psi(R)$ denote the integrand on the right side of* (4.7) *with $\mathbf{M} = J^{-1}$ and $B_i = \psi_i\psi_i^*\Lambda$.*

(a) For every polynomial p in the entries of $R_1, \dots, R_N$, the superfunction pF_ψ is integrable on $\mathcal{S}_{\Gamma,t}^N$ for all $t \in [0,1]$, and its integral does not depend on t. In particular,

$$\int_{\mathcal{S}_\flat^N} pF_\psi \,\mathrm{d}R = \int_{(\mathcal{S}+\mathrm{i}\eta_0\Gamma)^N} pF_\psi \,\mathrm{d}R.$$

(b) For $t \in [0,1]$ and $R_i \in \mathcal{S}_{\Gamma,t}$ with BB coordinates (P_i, X_i), the body of the BB part of the exponent of F_ψ satisfies

$$\mathrm{Re}\Big[-\sum_{i,j}\frac{(J^{-1})_{ij}}{2}\,\mathrm{Tr}\,\big(R_i^{\mathrm{BB}}R_j^{\mathrm{BB}}\big) - \mathrm{i}\sum_i \zeta_i^*\Lambda_{\mathrm{B}}R_i^{\mathrm{BB}}\zeta_i\Big] \\ \le \sum_{i=1}^N\Big(-\frac{1}{2}\|P_i\|^2 - \frac{\eta}{4}\,\mathrm{Tr}\,f_0(X_i) + C_{k,\eta}\big(1+|\zeta_i|^4\big)\Big), \quad (4.14)$$

where $C_{k,\eta}$ depends only on k and η.

We prove Proposition 4.8 in Appendix B, and we now explain the two computations on which it rests. First, at each site, the real part in (4.14) is

$$-\frac{1}{2}\|P\|^2 - \frac{t(2-t)}{2}\|X\|^2 - (1-t)^2\eta\,\mathrm{Tr}\,f_0(X) \\ + (1-t)^2k(s_E^2+\eta^2) - \zeta^*\big((1-t)(f_0(X)-\eta)+X\big)\zeta.$$

Here the Gaussian term $-\frac{1}{2}t(2-t)\|X\|^2$ and the damping term $-(1-t)^2\eta\,\mathrm{Tr}\,f_0(X)$ together give decay along X for all t. The source term is bounded by $t\zeta^*f_0(X)\zeta + \eta|\zeta|^2$, since $f_0(X) + X > 0$. Second, the link terms do not grow along the deformation. The translation by η cancels in $R_{i+1} - R_i$, and the identity

$$\|X - X'\|^2 - \|f_0(X) - f_0(X')\|^2 = s_E^2\,\mathrm{Tr}\,\big(\mathrm{g}_X^2\mathrm{g}_{X'}^{-2} + \mathrm{g}_X^{-2}\mathrm{g}_{X'}^2 - 2\mathrm{Id}\big) \ge 0$$

shows that f_0 is a contraction for the Hilbert–Schmidt norm. It follows that the real part of the link terms is bounded by $-\frac{1}{2}W^2\|P_{i+1} - P_i\|^2$ on every $\mathcal{S}_{\Gamma,t}$. The supplementary script

`rep_cycle_check.py` checks the identities of Lemma 4.7 and the last identity at random points, and `rep_homotopy_check.py` checks the per-site identity above, the bound (4.14) for each site, and the bounds of Lemma B.3.

We now compare Proposition 4.8 with earlier results. Schäfer and Wegner [111] introduced the parametrization $P - \mathrm{i}\lambda TsT^{-1}$ of the bosonic Hubbard–Stratonovich variable, with T in the noncompact group, which makes the integral converge while keeping the integrand bounded (see the description in [60, p. 2]). The BB part of $\mathcal{S}$ has this form with $\lambda = s_E$, by Lemma 4.7(d). A rigorous version of their prescription was given by Zirnbauer [139]. Following [139, Section III.B], we say that an ensemble of Hermitian matrices is of *class C* if its matrices have the form $H = \left(\begin{smallmatrix} a & b \\ b^* & -a^T \end{smallmatrix}\right)$ with a Hermitian and b symmetric, so that $\mathrm{i}H$ lies in the Lie algebra of the compact symplectic group. For the Gaussian ensemble of class C, he shows in [139, Lemma 3.2 and Theorem 3.3] that the image of the Schäfer–Wegner parametrization is a manifold without boundary on which the Gaussian factor and the coupling to the vector variables are bounded, and he expresses the ensemble average of a generating function as a superintegral over a single supermatrix. The class of the GUE, with the noncompact group $U(n_A, n_R)$, is treated in the same way in [139, Section IV.G]. The corresponding identity for a different domain, due to Pruisken and Schäfer, was established by Fyodorov for $U(p, q)$ [52] (see [60, p. 4]), by Fyodorov, Wei, and Zirnbauer for $O(p, q)$ [60, Theorem 1], where the integral becomes a sum of integrals over subdomains with alternating signs, and by Müller-Hill and Zirnbauer for general noncompact groups [96, Theorem 2.1]. In [96, Appendix B], they also sketch an explicit deformation of the Schäfer–Wegner domain into the flat domain, for a single matrix variable and without odd variables, along a homotopy of a different form.

All these results concern a single matrix variable. In [139], the convergence in the noncompact directions comes from the coupling to the vector variables, and [60, 96] assume that the source A satisfies $As > 0$, in the notation of [60]. (In the chains of [117], the commuting variables are bosonized instead, as explained at the beginning of this section.) In contrast, Proposition 4.8 concerns N matrix variables coupled by the nearest-neighbor precision (4.8), with a source $\zeta_i \zeta_i^* \Lambda_{\mathrm{B}}$ at each site whose product with Λ_{B} has rank one. The source alone then produces decay in only some of the noncompact directions, and the decay must come from the damping η instead. What is new in Appendix B is the control of the link terms, through the contraction property of f_0, and an explicit deformation onto the translated cycle with a majorant that is uniform along the deformation. The argument is elementary, and we give it in full.

4.5 The open chain

We now define the objects of the chain and state the exact representation. Recall the cycle $\mathcal{S}$ of Definition 4.6. The superfunctions on $\mathcal{S}$ below are even, with values in the

Grassmann algebra generated by the odd entries of their arguments, and integrals over $\mathcal{S}$ are taken in the sense of Definition 4.6.

Definition 4.9. Let Γ be a source supermatrix.

(a) For $S \in \mathcal{S}$ and a $(2k|2k)$ supermatrix $Y = \operatorname{diag}(Y^{\mathrm{B}} \mid Y^{\mathrm{F}})$, we define the *onsite weight* and the *marked onsite weight*

$$\begin{aligned} w_\Gamma(S) &= e^{-\operatorname{STr}(S^2)/2 - \mathrm{i}\eta_0 \operatorname{STr}(S\Gamma)} \operatorname{SDet}(E - S)^{-1}, \\ w_{\Gamma,Y}(S) &= e^{-\operatorname{STr}(S^2)/2 - \mathrm{i}\eta_0 \operatorname{STr}(S\Gamma)} \operatorname{SDet}(E - S + Y)^{-1}, \end{aligned} \tag{4.15}$$

the second for Y in a neighborhood of zero that may depend on S. We let w_0 denote w_Γ with $\Gamma = 0$, so that $w_\Gamma = w_0 e^{-\mathrm{i}\eta_0 \operatorname{STr}(S\Gamma)}$.

(b) For $t > 0$ and $S, S' \in \mathcal{S}$, we define the *heat kernel* $k_t(S, S') = e^{-\operatorname{STr}(S-S')^2/(4t)}$, and we let K_t denote the operator

$$(K_t f)(S) = \int_{\mathcal{S}} k_t(S, S') f(S') \, \mathrm{d}S',$$

where the Berezin integral is over the odd entries of S'. We set $\tau_{\mathrm{phys}} = (2W^2)^{-1}$, so that $k_{\tau_{\mathrm{phys}}}(S, S') = e^{-W^2 \operatorname{STr}(S-S')^2/2}$ is the link factor of (4.10).

(c) We define the *transfer operator* $T_\Gamma = M_{w_\Gamma} K_{\tau_{\mathrm{phys}}}$, where M_w denotes multiplication by w, so that

$$(T_\Gamma f)(S) = w_\Gamma(S) \int_{\mathcal{S}} k_{\tau_{\mathrm{phys}}}(S, S') f(S') \, \mathrm{d}S', \tag{4.16}$$

so that the heat link acts first and the onsite weight is applied after it. We define its *analytic transpose* $T_\Gamma^{\mathrm{tr}} = K_{\tau_{\mathrm{phys}}} M_{w_\Gamma}$, which applies the weight first. For integers $m \geq 1$, we define the *partial chains*

$$f_m = T_\Gamma^m 1, \qquad \ell_m = (T_\Gamma^{\mathrm{tr}})^m 1.$$

We say that ℓ_m, and more generally a product of steps T_Γ^{tr} with insertions, is built in the *analytic transpose order*.

(d) For superfunctions f and g on $\mathcal{S}$ such that fg is integrable on $\mathcal{S}$, we define the bilinear pairing $\mathcal{B}_{\mathcal{S}}(f, g) = \int_{\mathcal{S}} f(S) g(S) \, \mathrm{d}S$.

(e) For a superfunction F on $\mathcal{S}^N$, we define the *chain integral*

$$\mathcal{Z}_N(\Gamma; F) = \int_{\mathcal{S}^N} F(S_1, \ldots, S_N) \prod_{i=1}^{N} w_\Gamma(S_i) \prod_{i=1}^{N-1} k_{\tau_{\mathrm{phys}}}(S_i, S_{i+1}) \, \mathrm{d}S, \tag{4.17}$$

whenever the integrand is integrable, and we write $\mathcal{Z}_N(\Gamma) = \mathcal{Z}_N(\Gamma; 1)$ for the chain integral of the open chain without insertion.

The superdeterminant in (4.15) is defined on all of $\mathcal{S}$. Indeed, let A, B, C, D denote the blocks of $E - S$. By Lemma 4.7(b), the block A is invertible. Using Lemma 3.11(a), we obtain

$$\mathrm{SDet}(E-S)^{-1} = \frac{\det(D - CA^{-1}B)}{\det A}.$$

The right side is a polynomial in the FF and odd entries of S and in the entries of A^{-1}, divided by $\det A$. The same holds for $w_{\Gamma,Y}$ when $\|Y^{\mathrm{B}}\|_{\mathrm{op}}$ is smaller than s_E times the smallest eigenvalue of g_X^2, and the derivatives of $w_{\Gamma,Y}$ in Y at $Y = 0$ are defined at every point of $\mathcal{S}$. We will need three facts about the heat kernel, which we prove in Appendix B.

Lemma 4.10. *Let $t, t' > 0$. For all $S, S'' \in \mathcal{S}$, the superfunctions $k_t(S, \cdot)$ and $k_t(S, \cdot)k_{t'}(\cdot, S'')$ are integrable on $\mathcal{S}$, and*

$$(K_t 1)(S) = 1, \qquad \int_{\mathcal{S}} k_t(S, S')k_{t'}(S', S'')\,\mathrm{d}S' = k_{t+t'}(S, S'').$$

Moreover, $k_t(S, S') = k_t(S', S)$, and the body of $k_t(S, S')$ has modulus at most $e^{-(\|P-P'\|^2+\|M-M'\|^2)/(4t)}$, where (P, X, M) and (P', X', M') are the coordinates of S and S'.

Since $\mathrm{d}S$ is normalized by (4.6), the identity $K_t 1 = 1$ holds without a constant factor, for all $t > 0$ and all W; the reason is the cancellation of determinants after (4.7), with $N = 1$ and $\mathbf{M} = (2t)^{-1}$.

We can now state the exact representation.

Proposition 4.11. *Let Γ be a source supermatrix.*

(a) The integrand of $\mathcal{Z}_N(\Gamma)$ is integrable on $\mathcal{S}^N$, and

$$T_N^{(k)}(\Gamma) = e^{N\eta_0^2\,\mathrm{STr}(\Gamma^2)/2}\,\mathcal{Z}_N(\Gamma). \tag{4.18}$$

(b) The partial chains f_m and ℓ_m are defined by absolutely convergent integrals, $\ell_m = K_{\tau_{\mathrm{phys}}} f_m$, and for all integers $1 \le m \le N/2$,

$$\mathcal{Z}_N(\Gamma) = \mathcal{B}_{\mathcal{S}}(1, T_\Gamma^N 1) = \mathcal{B}_{\mathcal{S}}\big(\ell_m, T_\Gamma^{N-2m} f_m\big). \tag{4.19}$$

Moreover, $\mathcal{B}_{\mathcal{S}}(T_\Gamma^{\mathrm{tr}} f, g) = \mathcal{B}_{\mathcal{S}}(f, T_\Gamma g)$ whenever both sides are given by absolutely convergent iterated integrals.

(c) Let Y be local sources. Then $T_N^{(k)}(\Gamma; Y)$ is holomorphic in the entries of Y for $\|Y^{\mathrm{B}}\|_{\mathrm{op}} < \eta$, and for every multi-index $\alpha = (\alpha_1, \dots, \alpha_N)$, where α_i is a multi-index in the entries of Y_i,

$$\partial_Y^\alpha T_N^{(k)}(\Gamma; Y)\big|_{Y=0} = e^{N\eta_0^2\,\mathrm{STr}(\Gamma^2)/2} \int_{\mathcal{S}^N} \prod_{i=1}^{N} \big(\partial_{Y_i}^{\alpha_i} w_{\Gamma,Y_i}(S_i)\big|_{Y_i=0}\big) \prod_{i=1}^{N-1} k_{\tau_{\mathrm{phys}}}(S_i, S_{i+1})\,\mathrm{d}S, \tag{4.20}$$

with an integrable integrand. If $Y^{\mathrm{B}} = 0$, then $T_N^{(k)}(\Gamma; Y)$ equals the right side of (4.20) with $\alpha = 0$ and $w_{\Gamma,Y_i}(S_i)$ in place of the derivatives, for all FF blocks Y_i^{F}.

In particular,

$$\mathcal{Z}_N(\Gamma_0) = 1, \qquad T_N^{(k)}(\Gamma) = e^{N\eta_0^2 \,\mathrm{STr}(\Gamma^2)/2}\, \frac{\mathcal{Z}_N(\Gamma)}{\mathcal{Z}_N(\Gamma_0)}. \tag{4.21}$$

Later sections use only the ratio form in (4.21), which holds for every normalization of $\mathrm{d}S$, and its prefactor tends to one together with its derivatives in the sources, since $N\eta_0^2 = (N\rho(E)^2)^{-1}$.

Proof. We begin with (c), which contains (a) as the case $\alpha = 0$, $Y = 0$. Holomorphy was shown in Lemma 4.3. In (4.5), the local sources appear only through the factor $e^{\mathrm{i}\sum_i \psi_i^* \Lambda Y_i \psi_i}$. For $\|Y^{\mathrm{B}}\|_{\mathrm{op}} \le \eta/2$, the body of the integrand is bounded by $e^{-\eta|\zeta|^2/2}$ times polynomials in ζ, uniformly in Y, so we may differentiate under the integral. We obtain

$$\partial_Y^\alpha T_N^{(k)}(\Gamma;Y)\big|_{Y=0} = \int p_\alpha(\psi) \exp\Big(\mathrm{i}\sum_i \psi_i^* \Lambda(E+\mathrm{i}\eta_0\Gamma)\psi_i - \frac{1}{2}\sum_{i,j} J_{ij}\,\mathrm{STr}\,\big(\psi_i\psi_i^*\Lambda\,\psi_j\psi_j^*\Lambda\big)\Big)\,\mathrm{d}\psi, \tag{4.22}$$

where

$$p_\alpha(\psi) = \partial_Y^\alpha e^{\mathrm{i}\sum_i \psi_i^* \Lambda Y_i \psi_i}\big|_{Y=0}$$

is a polynomial in ψ.

For fixed ψ, we first apply (4.7) with $\mathbf{M} = J^{-1}$ and $B_i = \psi_i\psi_i^*\Lambda$, and we then deform the cycle by Proposition 4.8(a) with $p = 1$. Using (4.4), we find

$$\partial_Y^\alpha T_N^{(k)}(\Gamma;Y)\big|_{Y=0} = \int \mathrm{d}\psi \int_{(\mathcal{S}+\mathrm{i}\eta_0\Gamma)^N} p_\alpha(\psi)\, e^{\mathrm{i}\sum_i \psi_i^*\Lambda(E+\mathrm{i}\eta_0\Gamma-R_i)\psi_i}\, e^{-\frac{1}{2}\sum_{i,j}(J^{-1})_{ij}\,\mathrm{STr}(R_iR_j)}\,\mathrm{d}R. \tag{4.23}$$

We next show that the integrand in (4.23) is integrable on the product, so that the two integrals may be exchanged. Let $R_i = S_i + \mathrm{i}\eta_0\Gamma$ with $S_i \in \mathcal{S}$. The body of the BB part of the ψ_i-exponent has real part

$$-\zeta_i^*\,\mathrm{Im}\,\big(\Lambda_{\mathrm{B}}(E - S_i^{\mathrm{BB}})\big)\zeta_i = -s_E\,\zeta_i^*\,\mathrm{g}_{X_i}^2\zeta_i,$$

by Lemma 4.7(b). The coefficients of the integrand are bounded by polynomials in ζ, in the coordinates of the S_i and in the entries of $\eta_0\Gamma$, times

$$\exp\Big(-s_E\sum_i \zeta_i^*\,\mathrm{g}_{X_i}^2\zeta_i\Big)\,\Big|e^{-\frac{1}{2}\sum_{i,j}(J^{-1})_{ij}\,\mathrm{STr}(R_iR_j)}\Big|_{\varnothing},$$

where $|\cdot|_\varnothing$ denotes the modulus of the body. For fixed X_i, the integral of the first factor times a polynomial of degree $2q$ in ζ_i is at most

$$C_q s_E^{-q}\|\mathrm{g}_{X_i}^{-2}\|_{\mathrm{op}}^q \det(s_E\mathrm{g}_{X_i}^2)^{-1} \le C_q s_E^{-2q-2k}\big(s_E + 2\|X_i\|_{\mathrm{op}}\big)^q,$$

where we used Lemma 4.7.

We now bound the second factor. The case $t = 0$, $\zeta = 0$ of (4.14) implies that its BB part is bounded by

$$\prod_{i=1}^{N} \exp\Big(-\frac{1}{2}\|P_i\|^2 - \frac{\eta}{4} \operatorname{Tr} f_0(X_i) + C_{k,\eta} \Big).$$

Its FF part is a Gaussian in the M_i times $e^{C|M|}$, since $J^{-1} \geq \mathrm{Id}$. The polynomial factors in X_i of the previous paragraph are dominated by $e^{-\eta \operatorname{Tr} f_0(X_i)/4}$, since $\operatorname{Tr} f_0(X_i) \geq \|X_i\|$. Altogether, the integrand is integrable on the product.

Since the supervectors ψ_i have $4kN$ odd coordinates in all and the supermatrices R_i have $8k^2N$, both even numbers, Lemma 3.6(f) shows that the two Berezin integrals may be taken in either order without a sign, and Fubini's theorem allows us to integrate over ψ first.

We now integrate over ψ. At the site i, Lemma 4.7(b) yields

$$\operatorname{Re}\big[-\mathrm{i}\Lambda_{\mathrm{B}}\big(E + \mathrm{i}\eta_0\Gamma + Y_i - R_i\big)^{\mathrm{BB}}\big] = s_E \mathrm{g}_{X_i}^2 + \operatorname{Im}\big(\Lambda_{\mathrm{B}} Y_i^{\mathrm{B}}\big),$$

which is positive definite for Y_i near zero. Applying Proposition 3.12 and differentiating under the ψ-integral, we obtain

$$\int p_\alpha(\psi)\, e^{\mathrm{i}\sum_i \psi_i^* \Lambda(E+\mathrm{i}\eta_0\Gamma - R_i)\psi_i}\, \mathrm{d}\psi = \prod_{i=1}^{N} \partial_{Y_i}^{\alpha_i} \operatorname{SDet}\big(E + \mathrm{i}\eta_0\Gamma + Y_i - R_i\big)^{-1}\big|_{Y_i=0}.$$

It remains to substitute $R_i = S_i + \mathrm{i}\eta_0\Gamma$, under which $\int_{(\mathcal{S}+\mathrm{i}\eta_0\Gamma)^N} \mathrm{d}R$ becomes $\int_{\mathcal{S}^N} \mathrm{d}S$ by Definition 4.6. The links are unchanged by (4.8), and

$$\frac{1}{2}\sum_{i=1}^{N} \operatorname{STr}(S_i + \mathrm{i}\eta_0\Gamma)^2 = \frac{1}{2}\sum_{i=1}^{N} \operatorname{STr} S_i^2 + \mathrm{i}\eta_0 \sum_{i=1}^{N} \operatorname{STr}(S_i\Gamma) - \frac{N\eta_0^2}{2} \operatorname{STr}\Gamma^2,$$

where we used $\operatorname{STr}(S_i\Gamma) = \operatorname{STr}(\Gamma S_i)$. Collecting the factors yields (4.20), whose integrand is integrable by the bounds above.

Now let $Y^{\mathrm{B}} = 0$. The FF block of $E + \mathrm{i}\eta_0\Gamma + Y - H$ appears in the superdeterminant only through the determinant of a Schur complement. Each side of the claimed identity is therefore a polynomial in the entries of the Y_i^{F}, and their Taylor coefficients at zero agree by (4.20). This proves (c) and (a).

We next prove (b). Recall that $K_{\tau_{\mathrm{phys}}} 1 = 1$ by Lemma 4.10, so that $T_\Gamma 1 = w_\Gamma$. Using (4.16) and induction, we find that

$$f_m(S_1) = w_\Gamma(S_1) \int_{\mathcal{S}^{m-1}} \prod_{i=2}^{m} w_\Gamma(S_i) \prod_{i=1}^{m-1} k_{\tau_{\mathrm{phys}}}(S_i, S_{i+1})\, \mathrm{d}S_2 \cdots \mathrm{d}S_m$$

has m weights and $m-1$ links. Moreover,

$$(T_\Gamma^{\mathrm{tr}})^m 1 = K_{\tau_{\mathrm{phys}}} (M_{w_\Gamma} K_{\tau_{\mathrm{phys}}})^{m-1} w_\Gamma = K_{\tau_{\mathrm{phys}}} T_\Gamma^m 1,$$

which shows that $\ell_m = K_{\tau_{\text{phys}}} f_m$ has one more link.

The integrand of each of these iterated integrals is a factor of the integrand of $\mathcal{Z}_N(\Gamma)$, possibly multiplied by one more link, whose body has modulus at most one by Lemma 4.10. The bounds in the proof of (a), with $\zeta = 0$, show that it is integrable for each value of the free variable. The function $T_\Gamma^{N-2m} f_m$ has $N - m$ weights and $N - m - 1$ links, so its pairing with ℓ_m has N weights and $N - 1$ links, arranged as in Figure 4.1. Each site carries $8k^2$ odd coordinates, an even number, so that the Berezin integrals over different sites may be taken in any order, and Fubini's theorem identifies both pairings in (4.19) with (4.17) for $F = 1$.

The transpose identity follows from Fubini's theorem and the symmetry $k_{\tau_{\text{phys}}}(S, S') = k_{\tau_{\text{phys}}}(S', S)$, without signs, since all superfunctions involved are even. Finally, recall from Lemma 4.3 that $T_N^{(k)}(\Gamma_0) = 1$ and $\operatorname{STr} \Gamma_0^2 = 0$. Inserting these into (4.18) at $\Gamma = \Gamma_0$, we conclude that $\mathcal{Z}_N(\Gamma_0) = 1$, and (4.21) follows. □

4.6 The saddle manifold

We now describe the critical points of the onsite weight, around which Part IV analyzes the chain. Recall the onsite weight w_Γ of (4.15), and write $w_\Gamma(S) = e^{-\mathcal{F}(S) - \mathrm{i}\eta_0 \operatorname{STr}(S\Gamma)}$ with $\mathcal{F}(S) = \frac{1}{2} \operatorname{STr} S^2 + \operatorname{STr} \log(E - S)$. The critical points of $\mathcal{F}$ solve $S = (E - S)^{-1}$, or equivalently $S^2 - ES + 1 = 0$. Let Q be an even supermatrix of size $(2k|2k)$ with $Q^2 = \mathrm{Id}$, and let $S = \kappa_E - \mathrm{i}s_E Q$. Then $E - S = \kappa_E + \mathrm{i}s_E Q$, and $(\kappa_E + \mathrm{i}s_E Q)(\kappa_E - \mathrm{i}s_E Q) = \kappa_E^2 + s_E^2 = 1$, so that

$$(E - S)^{-1} = \kappa_E - \mathrm{i}s_E Q = S. \tag{4.24}$$

Every such supermatrix S is therefore a critical point of $\mathcal{F}$.

For $p \in \{0, \dots, 2k\}$, let $R_*^{(p)}$ denote the diagonal supermatrix with BB block $\kappa_E - \mathrm{i}s_E \Lambda_{\mathrm{B}}$ and FF block $\kappa_E - \mathrm{i}s_E \operatorname{diag}(\mathrm{Id}_p, -\mathrm{Id}_{2k-p})$, which is of this form. The supermatrices $V R_*^{(p)} V^{-1}$, where V is an even invertible supermatrix of size $(2k|2k)$ with entries in a Grassmann algebra, form $2k + 1$ complex orbits, indexed by the *signature* p, and we call their union the *saddle manifold* (see Section 16.3, where p is fixed and $R_*^{(p)}$ is written R_*). The orbits of signature $p \neq k$ have more commuting than anticommuting normal directions, and their contributions to the chain become exponentially small in W after a macroscopic number of steps, as explained in Section 2.4 and proved in Lemma 20.2. The limit of the chain is therefore carried by the complex orbit of $\kappa_E - \mathrm{i}s_E \Lambda$, the orbit of signature $p = k$.

We call the *balanced component* of the saddle manifold the real form of this orbit whose body is the set of matrices

$$\kappa_E - \mathrm{i}s_E \operatorname{diag}\left(g \Lambda_{\mathrm{B}} g^{-1} \mid u \Lambda_{\mathrm{F}} u^{-1}\right),$$

where g ranges over the complex $2k \times 2k$ matrices with $g^* \Lambda_{\mathrm{B}} g = \Lambda_{\mathrm{B}}$ and u over the unitary $2k \times 2k$ matrices. Here the body of the orbit is the complex manifold of the matrices of this form with $g, u \in \mathrm{GL}(2k, \mathbb{C})$, since the body of $V R_*^{(k)} V^{-1}$ is the conjugate of $R_*^{(k)}$ by the body of V. By a *real form* of a complex manifold we mean, as for $\mathbb{R}^n$ in $\mathbb{C}^n$ or the unitary group in $\mathrm{GL}(n, \mathbb{C})$, a real submanifold whose tangent space at each point spans the complex tangent space over $\mathbb{C}$ and has half its real dimension. The matrices with the above restrictions on g and u form such a submanifold. At the point $\kappa_E - \mathrm{i} s_E \Lambda$, for instance, the tangent space of the orbit consists of the matrices $-\mathrm{i} s_E \operatorname{diag}(B \mid C)$ with B and C complex and block off-diagonal, and that of the submanifold of those with B anti-Hermitian and C Hermitian; at the other points, this follows by conjugation.

The cycle $\mathcal{S}$ passes through this body, by Lemma 4.7(d). In Section 5, we describe the balanced component, together with its odd directions, as a homogeneous superspace: its points are the supermatrices $\kappa_E - \mathrm{i} s_E Q$, where Q is the field on the target space $\mathcal{M}_k$ of the sigma model, and the odd elements of $\mathfrak{gl}(2k|2k)$ act on it infinitesimally. Since the supertrace and the superdeterminant are invariant under conjugation, by Lemma 3.11, the function $\mathcal{F}$ is constant on each component of the saddle manifold, and on the balanced component the onsite weight varies only through the small source term.

The links $e^{-W^2 \operatorname{STr}(S_{i+1} - S_i)^2/2}$ force neighboring variables to be close. Heuristically, the chain therefore stays near the saddle manifold, its normal fluctuations relax over about W sites, and its tangent variables on the balanced component perform a random walk whose scaling limit is the heat semigroup on $\mathcal{M}_k$ of Section 5. This heuristic is explained in Sections 2.3 and 5.1, and Part IV makes it rigorous.

4.7 Symmetry and the source factor

We record the symmetries of the weights and the size of the source factor, which Part IV uses in the tube around the saddle manifold.

Remark 4.12. We let $\mathrm{GL}(2k|2k)$ denote the group of even invertible $(2k|2k)$ supermatrices whose entries lie in a Grassmann algebra, and for an even supermatrix A we let the stabilizer of A be the subgroup of $V \in \mathrm{GL}(2k|2k)$ with $VA = AV$. For $V \in \mathrm{GL}(2k|2k)$ and even supermatrices S and S', Lemma 3.11(b) and (c) give

$$\begin{aligned} \operatorname{STr}(VSV^{-1})^2 &= \operatorname{STr} S^2, \\ \operatorname{SDet}\left(E - VSV^{-1}\right) &= \operatorname{SDet}(E - S), \\ \operatorname{STr}\left(V(S - S')V^{-1}\right)^2 &= \operatorname{STr}(S - S')^2. \end{aligned}$$

It follows that $w_0(VSV^{-1}) = w_0(S)$ and $k_t(VSV^{-1}, VS'V^{-1}) = k_t(S, S')$, wherever both sides are defined, and $\operatorname{STr}(VSV^{-1}\Gamma) = \operatorname{STr}(S\Gamma)$ if V commutes with Γ. By (4.24), the critical points of w_0 include the supermatrices $\kappa_E - \mathrm{i} s_E Q$ with $Q^2 = \mathrm{Id}$, and the stabilizer

of $\kappa_E - \mathrm{i}s_E\Lambda$ consists of the $V \in \mathrm{GL}(2k|2k)$ that commute with Λ. These are algebraic identities. Conjugation does not preserve the cycle $\mathcal{S}$, and invariance statements for integrals over $\mathcal{S}$ need a separate argument. Conjugation by an element of $\mathrm{GL}(2k|2k)$ with body $\mathrm{diag}(g \mid u)$, where $g^*\Lambda_{\mathrm{B}}g = \Lambda_{\mathrm{B}}$ and u is unitary, preserves the body of the balanced component, defined in Section 4.6, and such an element fixes the point $\kappa_E - \mathrm{i}s_E\Lambda$ if and only if it commutes with Λ.

The BB part of the source factor $w_\Gamma / w_0 = e^{-\mathrm{i}\eta_0 \operatorname{STr}(S\Gamma)}$ is real and at most one on the body of the balanced component of the saddle manifold. Part IV also uses bounds on its derivatives there. Since these are stated in the geometry of Section 5, we give them in Lemma B.5 of Appendix B.8.

4.8 The anchored chain

In Part IV, we need a priori bounds on the partial chains $\ell_n = K_{\tau_{\mathrm{phys}}} T_\Gamma^n 1$, as functions of their free variable (see Section 15). The partial chain ℓ_n is a chain of n sites whose first site is linked to a fixed point S of the cycle, and we say that it is *anchored* at S. The Gaussian part of a partial chain can be integrated as in Section 4.3, read backward. The result is an average over an $n \times n$ Gaussian matrix of an inverse superdeterminant in which S appears linearly. To state it, we let Δ_n denote the Neumann Laplacian on $[n]$, and we set

$$J_n = \big(\mathrm{Id}_n - W^2\Delta_n + W^2 e_1 e_1^*\big)^{-1}, \qquad \varsigma = W^2 J_n e_1 \in \mathbb{R}^n, \qquad \sigma_n = \sum_{i=1}^n \varsigma_i. \tag{4.25}$$

The matrix J_n^{-1} is the precision of the chain on $[n]$ with an additional link of weight W^2 from the site 1 to the fixed point. It has positive diagonal, nonpositive off-diagonal entries, and positive row sums, so that all entries of J_n are positive. Since $J_n^{-1}\mathbf{1} = \mathbf{1} + W^2 e_1$, we have

$$J_n\mathbf{1} = \mathbf{1} - \varsigma, \qquad \sigma_n = W^2(1 - \varsigma_1), \qquad \mathbf{1}^T J_n \mathbf{1} = n - \sigma_n, \qquad W^2 - W^4 (J_n)_{11} = \sigma_n, \tag{4.26}$$

and in particular $0 < \varsigma_i < 1$ for all i. We let $H^{(n)}$ denote the $n \times n$ complex Hermitian Gaussian matrix with $\mathbb{E} H_{ij}^{(n)} H_{kl}^{(n)} = \delta_{il}\delta_{jk}(J_n)_{ij}$, defined as in Definition 1.1.

Proposition 4.13. *Let Γ be a source supermatrix, let $n \geq 1$, and set $\ell_n = K_{\tau_{\mathrm{phys}}} T_\Gamma^n 1$. Then for all $S \in \mathcal{S}$,*

$$\ell_n(S) = \exp\Big(-\frac{\sigma_n}{2} \operatorname{STr} S^2 - \mathrm{i}\eta_0\sigma_n \operatorname{STr}(S\Gamma) - \frac{\eta_0^2(n - \sigma_n)}{2} \operatorname{STr}\Gamma^2 \Big) \\ \times \mathbb{E} \operatorname{SDet}\Big(E - H^{(n)} - \varsigma \otimes S + \mathrm{i}\eta_0 (\mathbf{1} - \varsigma) \otimes \Gamma \Big)^{-1}, \tag{4.27}$$

where $\varsigma \otimes S$ denotes the family $(\varsigma_i S)_{i\in[n]}$, in the notation of Section 4.1. More generally, let $\ell_n^{(\alpha)}$ denote the function defined by the same iterated integral as ℓ_n, with the weights $w_\Gamma(S_i)$ at the sites of the chain replaced by the derivatives $\partial_{Y_i}^{\alpha_i} w_{\Gamma,Y_i}(S_i)|_{Y_i=0}$, where S_1 is the site linked to S. Then $\ell_n^{(\alpha)}(S)$ is the derivative ∂_Y^α at $Y = 0$ of the right side of (4.27) *with Y_i added to the ith diagonal block of the supermatrix in the superdeterminant.*

The supermatrix in (4.27) has the blocks

$$\delta_{ij}\big(E - \varsigma_i S + \mathrm{i}\eta_0(1-\varsigma_i)\Gamma\big) - H_{ij}^{(n)}\mathrm{Id}, \qquad i,j \in [n].$$

The BB block of this supermatrix is invertible. Indeed, after multiplication by $\mathrm{Id}_n \otimes \Lambda_{\mathrm{B}}$, its body has imaginary part $\mathrm{diag}_i(\varsigma_i s_E \mathrm{g}_X^2 + (1-\varsigma_i)\eta) > 0$ by Lemma 4.7(b), where X is the coordinate of S. The same argument as in Lemma 4.3 shows that the expectation is finite.

Proof. Recall from Definition 4.9 that $\ell_n(S)$ is the integral over $(S_1, \dots, S_n) \in \mathcal{S}^n$ of

$$k_{\tau_{\mathrm{phys}}}(S, S_1) \prod_{i=1}^{n} w_\Gamma(S_i) \prod_{i=1}^{n-1} k_{\tau_{\mathrm{phys}}}(S_i, S_{i+1}).$$

We first write the superdeterminants as Gaussian integrals. Using Lemma 4.7(b) and Proposition 3.12, we have

$$\mathrm{SDet}(E - S_i)^{-1} = \int e^{\mathrm{i}\psi_i^* \Lambda(E-S_i)\psi_i}\, \mathrm{d}\psi_i$$

on $\mathcal{S}$. The bounds in the proof of Proposition 4.11, with the damping $e^{-\eta \operatorname{Tr} f_0(X_i)}$ of the weights in place of (4.14), show that the resulting integrand is integrable on $\mathcal{S}^n$ times the space of the ψ_i. As in the proof of Proposition 4.11, the numbers of odd coordinates involved are even, and using Fubini's theorem, we obtain

$$\ell_n(S) = \int \ell_n(S;\psi)\, \mathrm{d}\psi,$$

where

$$\ell_n(S;\psi) = \int_{\mathcal{S}^n} \exp\Big(-\frac{1}{2}\sum_{i,j=1}^{n} (J_n^{-1})_{ij} \operatorname{STr}(S_i S_j) + \sum_{i=1}^{n} \operatorname{STr}(S_i b_i) - \frac{W^2}{2}\operatorname{STr} S^2 \Big)\, \mathrm{d}S_1 \cdots \mathrm{d}S_n \times e^{\mathrm{i}\sum_i \psi_i^* \Lambda E \psi_i}$$

and

$$b_i = W^2 \delta_{i1} S - \mathrm{i}\big(\psi_i \psi_i^* \Lambda + \eta_0 \Gamma\big).$$

Here we expanded $-\frac{W^2}{2}\operatorname{STr}(S_1 - S)^2$, and we used (4.4) and the analogue of (4.8) on $[n]$.

We next compute $\ell_n(S;\psi)$ for fixed ψ. The BB part of its exponent is of the form treated in Lemma B.4, with unit masses, the damping η coming from $-\mathrm{i}\eta_0 \operatorname{STr}(S_i\Gamma)$, and one link to the fixed point S. That lemma deforms each S_i from $\mathcal{S}$ onto the cycle with Hermitian BB block and FF block in $\kappa_E\mathrm{Id} + \mathrm{iHerm}(2k)$, which is a translate of $\mathcal{S}_\flat$ (see Appendix B.7). Completing the square on this cycle and applying Lemma B.2, we find

$$\ell_n(S;\psi) = \exp\Big(\frac{1}{2}\sum_{i,j}(J_n)_{ij}\operatorname{STr}(b_ib_j) - \frac{W^2}{2}\operatorname{STr}S^2\Big)e^{\mathrm{i}\sum_i \psi_i^*\Lambda E\psi_i}.$$

We now expand $\frac{1}{2}\sum_{i,j}(J_n)_{ij}\operatorname{STr}(b_ib_j)$. Together with the term $-\frac{1}{2}W^2\operatorname{STr}S^2$ of $\ell_n(S;\psi)$, the terms with S only, those with S and Γ, and those with Γ only give

$$\begin{aligned}
&\frac{W^4}{2}(J_n)_{11}\operatorname{STr}S^2 - \frac{W^2}{2}\operatorname{STr}S^2 = -\frac{\sigma_n}{2}\operatorname{STr}S^2,\\
&-\mathrm{i}\eta_0 W^2(J_n\mathbf{1})_1\operatorname{STr}(S\Gamma) = -\mathrm{i}\eta_0\sigma_n\operatorname{STr}(S\Gamma),\\
&-\frac{\eta_0^2}{2}\big(\mathbf{1}^TJ_n\mathbf{1}\big)\operatorname{STr}\Gamma^2 = -\frac{\eta_0^2(n-\sigma_n)}{2}\operatorname{STr}\Gamma^2,
\end{aligned}$$

where we used (4.26). Using $(J_n\mathbf{1})_i = 1-\varsigma_i$ and (4.4), we find that the remaining terms are

$$-\frac{1}{2}\sum_{i,j}(J_n)_{ij}\operatorname{STr}\big(\psi_i\psi_i^*\Lambda\psi_j\psi_j^*\Lambda\big) - \mathrm{i}\sum_i\varsigma_i\,\psi_i^*\Lambda S\psi_i - \eta_0\sum_i(1-\varsigma_i)\,\psi_i^*\Lambda\Gamma\psi_i.$$

The computation in the proof of Lemma 4.4, read backward with $H^{(n)}$ in place of H, shows that

$$\exp\Big(-\frac{1}{2}\sum_{i,j}(J_n)_{ij}\operatorname{STr}\big(\psi_i\psi_i^*\Lambda\psi_j\psi_j^*\Lambda\big)\Big) = \mathbb{E}\exp\Big(-\mathrm{i}\sum_{i,j}H_{ij}^{(n)}\psi_i^*\Lambda\psi_j\Big).$$

Altogether, $\ell_n(S;\psi)$ is the Gaussian prefactor in (4.27) times

$$\mathbb{E}\exp\Big(\mathrm{i}\sum_i\psi_i^*\Lambda\big(E-\varsigma_iS+\mathrm{i}\eta_0(1-\varsigma_i)\Gamma\big)\psi_i - \mathrm{i}\sum_{i,j}H_{ij}^{(n)}\psi_i^*\Lambda\psi_j\Big).$$

By Lemma 4.7(b), the body of the exponent has the real part

$$-\sum_{i=1}^n\zeta_i^*\big(\varsigma_i s_E\mathrm{g}_X^2 + (1-\varsigma_i)\eta\big)\zeta_i,$$

which is independent of $H^{(n)}$.

We may now exchange the expectation and the ψ-integral, and (4.27) follows from Proposition 3.12.

For the marked version, we argue as in the proof of Proposition 4.11(c). Differentiating under the integral at fixed $S_i \in \mathcal{S}$, we obtain

$$\partial_{Y_i}^{\alpha_i} \operatorname{SDet}(E - S_i + Y_i)^{-1}\big|_{Y_i=0} = \int p_{\alpha_i}(\psi_i)\, e^{\mathrm{i}\psi_i^* \Lambda (E-S_i)\psi_i}\, \mathrm{d}\psi_i,$$

with the polynomial p_{α_i} of (4.22). This polynomial does not involve $S_1, \dots, S_n$, and the argument above applies with the extra factor $\prod_i p_{\alpha_i}(\psi_i)$. With Y_i added to the ith diagonal block, the matrix $\varsigma_i s_E \mathrm{g}_X^2 + (1-\varsigma_i)\eta$ in the real part of the exponent above becomes $\varsigma_i s_E \mathrm{g}_X^2 + (1-\varsigma_i)\eta + \operatorname{Im}(\Lambda_{\mathrm{B}} Y_i^{\mathrm{B}})$, which is at least $((1-\varsigma_i)\eta - \|Y_i^{\mathrm{B}}\|_{\mathrm{op}})\mathrm{Id}$. For $\|Y_i^{\mathrm{B}}\|_{\mathrm{op}} \le (1-\varsigma_i)\eta/2$, the coefficients of the integrand are therefore bounded, uniformly in Y, by polynomials in ζ and in the entries of $H^{(n)}$ times $e^{-\sum_i (1-\varsigma_i)\eta|\zeta_i|^2/2}$. This allows us to differentiate under the ψ-integral, and the claim follows. □

4.9 Resolvent entries and insertions

We now state the dictionary between the resolvent $G(z) = (H-z)^{-1}$ and the chain. Derivatives of $T_N^{(k)}(\Gamma; Y)$ in the local sources are expectations of polynomials in resolvent entries. In the chain, they are integrals with polynomials in the entries of $(E - S_i)^{-1}$ inserted at the marked sites.

Let Γ be a source supermatrix with diagonal FF block. We let $z_1^{\mathrm{F}}, \dots, z_{2k}^{\mathrm{F}}$ denote the diagonal entries of $E + \mathrm{i}\eta_0\Gamma_{\mathrm{F}}$, and we set $z_a^{\mathrm{B}} = z_0$ for $a \le k$ and $z_a^{\mathrm{B}} = \bar{z}_0$ for $a > k$. We let $\mathbf{G}^{\mathrm{B}}$ and $\mathbf{G}^{\mathrm{F}}$ denote the $2kN \times 2kN$ matrices with entries

$$(\mathbf{G}^{\mathrm{B}})_{(i,a),(j,b)} = \delta_{ab}\, G_{ij}(z_a^{\mathrm{B}}), \qquad (\mathbf{G}^{\mathrm{F}})_{(i,a),(j,b)} = \delta_{ab}\, G_{ij}(z_a^{\mathrm{F}}),$$

indexed by sites $i, j \in [N]$ and flavors $a, b \in [2k]$. For local sources Y, we let $\mathbf{Y}^{\mathrm{B}}$ and $\mathbf{Y}^{\mathrm{F}}$ denote the block-diagonal matrices with (i,i) blocks Y_i^{B} and Y_i^{F}. We set $\mathcal{R}_N(\Gamma) = \operatorname{SDet}(E + \mathrm{i}\eta_0\Gamma - H)^{-1}$, so that $\mathbb{E}\mathcal{R}_N(\Gamma) = T_N^{(k)}(\Gamma)$, and by the proof of Lemma 4.3, $\mathcal{R}_N(\Gamma(u,b))$ is the product inside the expectation in (1.4). For $a \in [k]$, we set $e_a^+ = e_a + e_{k+a} \in \mathbb{C}^{2k}$ and $e_a^- = (e_a - e_{k+a})/(2\mathrm{i})$. For $S \in \mathcal{S}$ such that $E - S$ is invertible, we let $[(E-S)^{-1}]^{\mathrm{BB}}$ and $[(E-S)^{-1}]^{\mathrm{FF}}$ denote the BB and FF blocks of the inverse supermatrix, and we define

$$\gamma_a(S) = -(e_a^-)^T \big[(E-S)^{-1}\big]^{\mathrm{BB}} e_a^+. \tag{4.28}$$

Lemma 4.14. *Let Γ be a source supermatrix with diagonal FF block, and let Y be local sources with $\|Y^{\mathrm{B}}\|_{\mathrm{op}} < \eta$.*

(a) We have

$$T_N^{(k)}(\Gamma; Y) = \mathbb{E}\Big[\mathcal{R}_N(\Gamma)\, \frac{\det(\mathrm{Id} - \mathbf{G}^{\mathrm{F}}\mathbf{Y}^{\mathrm{F}})}{\det(\mathrm{Id} - \mathbf{G}^{\mathrm{B}}\mathbf{Y}^{\mathrm{B}})}\Big]. \tag{4.29}$$

(b) At every point of $\mathcal{S}$ where $E - S$ is invertible, which is all of $\mathcal{S}$ if $E \neq 0$ and a set of full measure if $E = 0$, we have

$$w_{\Gamma,Y}(S) = w_\Gamma(S)\ \mathrm{SDet}\left(\mathrm{Id} + (E-S)^{-1}Y\right)^{-1}. \tag{4.30}$$

If $Y = \mathrm{diag}(Y^{\mathrm{B}} \mid 0)$, the last factor equals $\det(\mathrm{Id} + [(E-S)^{-1}]^{\mathrm{BB}}Y^{\mathrm{B}})^{-1}$, and if $Y = \mathrm{diag}(0 \mid Y^{\mathrm{F}})$, it equals $\det(\mathrm{Id} + [(E-S)^{-1}]^{\mathrm{FF}}Y^{\mathrm{F}})$.

Consequently, in the sense of Taylor coefficients at $Y = 0$, the entries of $\mathbf{G}^{\mathrm{B}}$ and $\mathbf{G}^{\mathrm{F}}$ in (4.29) *correspond to the entries of $-[(E-S_i)^{-1}]^{\mathrm{BB}}$ and $-[(E-S_i)^{-1}]^{\mathrm{FF}}$ in* (4.20)*. In particular, let $\Gamma = \Gamma_0$, so that $z = z_0 = E + \mathrm{i}\eta$, and let $i_1, \dots, i_m \in [N]$ be distinct sites. Then, in the notation* (4.17)*,*

$$\mathbb{E}\,\mathrm{per}\left[G_{i_a i_b}(z)\right]_{a,b=1}^m = \mathcal{Z}_N\Big(\Gamma_0; \prod_{a=1}^m \big(-[(E-S_{i_a})^{-1}]_{11}^{\mathrm{BB}}\big)\Big), \tag{4.31}$$

$$\mathbb{E}\det\left[G_{i_a i_b}(z)\right]_{a,b=1}^m = \mathcal{Z}_N\Big(\Gamma_0; \prod_{a=1}^m \big(-[(E-S_{i_a})^{-1}]_{11}^{\mathrm{FF}}\big)\Big), \tag{4.32}$$

$$\mathbb{E}\,\mathrm{per}\left[\,\mathrm{Im}\,G_{i_a i_b}(z)\right]_{a,b=1}^m = \mathcal{Z}_N\Big(\Gamma_0; \prod_{a=1}^m \gamma_1(S_{i_a})\Big), \tag{4.33}$$

where per *denotes the permanent and* $\mathrm{Im}\,G = (G(z) - G(z)^*)/(2\mathrm{i})$*. If $m \le k$, then*

$$\mathbb{E}\prod_{a=1}^m \mathrm{Im}\,G_{i_a i_a}(z) = \mathcal{Z}_N\Big(\Gamma_0; \prod_{a=1}^m \gamma_a(S_{i_a})\Big). \tag{4.34}$$

Finally, for all $i \in [N]$ and integers $q \ge 1$, writing $y = [(E-S_i)^{-1}]^{\mathrm{BB}}$,

$$\mathbb{E}|G_{ii}(z)|^{2q} = \mathcal{Z}_N\Big(\Gamma_0; \sum_{r=0}^q \binom{q}{r}^2 (y_{11}\,y_{k+1,k+1})^{q-r}(y_{1,k+1}\,y_{k+1,1})^r\Big). \tag{4.35}$$

All chain integrals in (4.31)–(4.35) *have integrable integrands.*

Recall from Section 1.2 that $\ell_{i,N}^{\varepsilon}(0) = s_E^{-1}\,\mathrm{Im}\,G_{ii}(z_0)$; for a center $u \neq 0$, (4.34) applies with E replaced by $E + u/a_N$ and the same η. The local density $\ell_{i,N}^{\varepsilon}$, a random function attached to the site i, is unrelated to the partial chains ℓ_m of Definition 4.9, which are superfunctions on $\mathcal{S}$ built from m sites of the chain. The permanent (4.33) at distinct sites contains the products $\mathrm{Im}\,G_{i_a i_b}\,\mathrm{Im}\,G_{i_b i_a}$, and the product of diagonal entries is obtained from (4.34), which uses one pair of flavors for each site. The sites in (4.34) must be distinct. At a single site i, the two pairs of flavors 1 and 2 give, with $y = [(E-S_i)^{-1}]^{\mathrm{BB}}$, the insertion

$$\gamma_1(S_i)\gamma_2(S_i) + (e_1^-)^T y\, e_2^+\,(e_2^-)^T y\, e_1^+$$

for $\mathbb{E}(\mathrm{Im}\,G_{ii}(z))^2$, and not $\gamma_1(S_i)\gamma_2(S_i)$; moments at a single site are given by (4.35). In (4.35), the matrix side does not couple the retarded and advanced flavors, while the

insertion contains the entries $y_{1,k+1}$ and $y_{k+1,1}$, which couple them. Such terms distinguish coincident from distinct sites in Theorem 1.14.

Proof. We begin with (a). Let $D = E + \mathrm{i}\eta_0\Gamma - H$. The BB and FF blocks of D are block diagonal over the flavors, with (a,a) blocks $z_a^{\mathrm{B}} - H$ and $z_a^{\mathrm{F}} - H$ on $\mathbb{C}^N$. Using $(z-H)^{-1} = -G(z)$ and reordering the indices, we find

$$D^{-1} = -\operatorname{diag}\big(\mathbf{G}^{\mathrm{B}} \mid \mathbf{G}^{\mathrm{F}}\big), \qquad D + Y = D\big(\mathrm{Id} - \operatorname{diag}(\mathbf{G}^{\mathrm{B}} \mid \mathbf{G}^{\mathrm{F}})\, Y\big).$$

By Lemma 3.11(b), we now have

$$\operatorname{SDet}(D+Y)^{-1} = \operatorname{SDet}(D)^{-1} \frac{\det(\mathrm{Id} - \mathbf{G}^{\mathrm{F}}\mathbf{Y}^{\mathrm{F}})}{\det(\mathrm{Id} - \mathbf{G}^{\mathrm{B}}\mathbf{Y}^{\mathrm{B}})},$$

and taking expectations, we obtain (4.29).

For (b), the multiplicativity of SDet in Lemma 3.11(b) implies that

$$\operatorname{SDet}(E - S + Y) = \operatorname{SDet}(E-S)\operatorname{SDet}\big(\mathrm{Id} + (E-S)^{-1}Y\big),$$

and (4.30) follows from the definition (4.15). When Y has only a BB block, the supermatrix $\mathrm{Id} + (E-S)^{-1}Y$ has FF block Id and BF block zero. When Y has only an FF block, its BB block is Id and its FB block is zero. In these two cases, Lemma 3.11(a) and the definition of SDet give, respectively,

$$\operatorname{SDet}\big(\mathrm{Id} + (E-S)^{-1}Y\big) = \det\big(\mathrm{Id} + [(E-S)^{-1}]^{\mathrm{BB}}Y^{\mathrm{B}}\big),$$
$$\operatorname{SDet}\big(\mathrm{Id} + (E-S)^{-1}Y\big) = \det\big(\mathrm{Id} + [(E-S)^{-1}]^{\mathrm{FF}}Y^{\mathrm{F}}\big)^{-1}.$$

Note that the FF block $\kappa_E - \mathrm{i}M$ of $E - S$ is invertible for all M if $E \neq 0$, and outside the null set where M is singular if $E = 0$.

The correspondence follows by comparing (4.29), at $\Gamma = \Gamma_0$ where $\mathcal{R}_N(\Gamma_0) = 1$, with (4.20) and (4.30). We now verify the special cases. For (4.31), we set

$$Y_{i_a} = \operatorname{diag}(t_a E_{11} \mid 0), \qquad A = \sum_{a=1}^{m} t_a e_{i_a} e_{i_a}^*,$$

where E_{11} is the matrix unit of the first flavor. The letter E with two indices, such as E_{11} here or E_{ij} and E_{RR} in later sections, always denotes a matrix unit, and E without indices is the energy. Then (4.29) reduces to

$$T_N^{(k)}(\Gamma_0; Y) = \mathbb{E}\det\big(\mathrm{Id}_N - G(z)A\big)^{-1}.$$

The coefficient of $t_1 \cdots t_m$ in $\det(\mathrm{Id}_N - GA)^{-1}$ is the permanent of the submatrix of G with rows and columns $i_1, \ldots, i_m$, by the expansion of $\det(\mathrm{Id} - A)^{-1}$ into permanents (MacMahon's master theorem). For the chain side, (4.30) gives the factor

$$\prod_{a=1}^{m} \big(1 + t_a[(E - S_{i_a})^{-1}]_{11}^{\mathrm{BB}}\big)^{-1},$$

whose coefficient of $t_1 \cdots t_m$ is the insertion in (4.31).

For (4.32), we use the FF sources $Y_{i_a} = \operatorname{diag}(0 \mid t_a E_{11})$ instead. The coefficients of $t_1 \cdots t_m$ on the matrix side and on the chain side are then, respectively,

$$(-1)^m \mathbb{E} \det \big[G_{i_a i_b}(z)\big]_{a,b=1}^m \qquad \text{and} \qquad \prod_{a=1}^m \big[(E - S_{i_a})^{-1}\big]_{11}^{\mathrm{FF}}.$$

For (4.33), we use the sources

$$Y_{i_a} = \operatorname{diag}\big(t_a e_1^+ (e_1^-)^T \mid 0\big).$$

The matrix $e_1^+ (e_1^-)^T$ has rank one and $\mathbf{G}^{\mathrm{B}}$ is diagonal in the flavors, and therefore

$$\det\big(\mathrm{Id} - \mathbf{G}^{\mathrm{B}}\mathbf{Y}^{\mathrm{B}}\big) = \det\Big(\mathrm{Id}_N - A\,\frac{G(z) - G(\bar z)}{2\mathrm{i}}\Big) = \det\big(\mathrm{Id}_N - A \operatorname{Im} G(z)\big),$$

where we used $G(\bar z) = G(z)^*$. As for (4.31), the expansion into permanents shows that $\operatorname{per}[\operatorname{Im} G_{i_a i_b}]$ is the coefficient of $t_1 \cdots t_m$ in the inverse of this determinant. For the chain side, we set $y = [(E - S_{i_a})^{-1}]^{\mathrm{BB}}$, and we have

$$\det\big(\mathrm{Id} + t_a y e_1^+ (e_1^-)^T\big) = 1 + t_a (e_1^-)^T y e_1^+ = 1 - t_a \gamma_1(S_{i_a}),$$

and the inverses of these factors give the insertion.

For (4.34), we use the source $t_a e_a^+ (e_a^-)^T$ at the site i_a. Different sites now use different pairs of flavors. Since $\mathbf{G}^{\mathrm{B}}$ is diagonal in the flavors, the matrix side factors,

$$\det\big(\mathrm{Id} - \mathbf{G}^{\mathrm{B}}\mathbf{Y}^{\mathrm{B}}\big)^{-1} = \prod_{a=1}^m \big(1 - t_a \operatorname{Im} G_{i_a i_a}(z)\big)^{-1},$$

and the chain side is the product of the factors $(1 - t_a \gamma_a(S_{i_a}))^{-1}$, as above.

For (4.35), we use the source

$$Y_i = \operatorname{diag}\big(tE_{11} + t' E_{k+1,k+1} \mid 0\big).$$

On the matrix side, since $|G_{ii}(z)| \le \eta^{-1}$, we may expand both factors into geometric series for $|t|, |t'| < \eta$, and we obtain

$$\mathbb{E}\big(1 - tG_{ii}(z)\big)^{-1}\big(1 - t' G_{ii}(\bar z)\big)^{-1} = \sum_{q,q' \ge 0} t^q t'^{q'}\, \mathbb{E}\big[G_{ii}(z)^q G_{ii}(\bar z)^{q'}\big].$$

Since $G_{ii}(\bar z) = \overline{G_{ii}(z)}$, the coefficient of $t^q t'^q$ is the moment $\mathbb{E}|G_{ii}(z)|^{2q}$. The factor on the chain side is the inverse of

$$(1 + ty_{11})(1 + t' y_{k+1,k+1}) - tt' y_{1,k+1} y_{k+1,1},$$

and expanding first in its last term gives the coefficient in (4.35).

The insertions are Taylor coefficients of $\operatorname{SDet}(\mathrm{Id} + (E - S)^{-1}Y)^{-1}$, which are the derivatives of the marked weights in (4.20) divided by w_Γ. Their integrability follows from Proposition 4.11(c). □

The supplementary script `rep_dictionary_check.py` checks the matrix and chain sides of (4.31)–(4.35) separately, before the expectation, together with the coincident-site insertion above and the saddle values below, for $k = 2$ and a random Hermitian matrix of size five, to a relative error of order 10^{-12}.

The insertions take explicit values on the saddle manifold of Section 4.6. Let $S = \kappa_E - \mathrm{i}s_E Q$ with $Q^2 = \mathrm{Id}$. By (4.24), we have $-(E - S)^{-1} = -\kappa_E + \mathrm{i}s_E Q$, the supermatrix that Part III uses under the name orbit resolvent, in (7.16). At $Q = \Lambda$, its retarded and advanced diagonal entries $-\kappa_E \pm \mathrm{i}s_E$ are the values of $G_{ii}(E \pm \mathrm{i}0)$ for the semicircle law. Inserting (4.24) into (4.28), we find

$$\gamma_a(S) = \frac{s_E}{2}\big(Q_{aa} + Q_{a,k+a} - Q_{k+a,a} - Q_{k+a,k+a}\big)^{\mathrm{BB}},$$

which at $Q = \Lambda$ gives $\gamma_a(S) = s_E = \mathrm{Im}(-\kappa_E + \mathrm{i}s_E)$.

SECTION 5

The sigma model on a wire

This section defines the continuum object behind our main results, the partition function $\mathcal{Z}^{(k)}_{X,\varepsilon}(\mathbf{u},\mathbf{b})$ of Definition 1.3, and proves the general facts about it that the rest of the paper uses. The partition function is the Berezin integral of a heat semigroup on a symmetric superspace $\mathcal{M}_k$, the target space of the supersymmetric sigma model, and its generator is X times an invariant second-order operator, the Casimir operator, plus a potential that carries the sources. We explain heuristically in Section 5.1 why this object describes the chain of Section 4, and we make it precise in Definition 5.16. In Proposition 5.15, whose proof is in Appendix C, we construct the semigroup on weighted L^2 spaces and show that the partition function is well defined. In Lemma 5.19 and Corollary 5.20, we prove a localization identity of Parisi–Sourlas type and deduce that the partition function equals one without sources. The next section, Section 6, reduces the case of rank one to two scalar diffusions on the domain D of Definition 1.4, the radial and the singlet diffusion, develops the calculus of rank-one pairings that Part III uses, and shows in Lemma 6.15 how the singlet state $\Psi_t^{X,z}$ and the kernels K_X and I_X arise from the partition function.

Part III uses only these results and those of Section 6, together with Theorem 7.6, and the opening of Part II lists what Part IV takes from this section. From Section 5.2, the reader of Part III needs only the definitions of the field Q and of the invariant integral, the formula (5.14) for the Casimir operator C_k, the three properties of C_k in Lemma 5.8, and the integration by parts formula (5.13). The algebraic facts behind these definitions are contained in proofs and remarks, which this reader may skip, as well as Appendix C. After Definition 5.16 and Lemma 5.17, such a reader can go on to the statements of Section 5.4 and then to Section 6. At rank one, in Section 6, everything is explicit. The underlying manifold of $\mathcal{M}_1$ is a hyperbolic plane times a sphere, and on superfunctions that are invariant under the stabilizer of the base point, C_1 is an explicit second-order operator on the domain D (Lemma 6.8). We use the Grassmann calculus of Section 3 throughout, and we refer to Section 2 for the role of the sigma model in the proof.

5.1 Why this target space

We first explain heuristically why the limits of the source ratios $T_N^{(k)}$ should be described by a heat semigroup on a superspace of supermatrices Q with $Q^2 = \mathrm{Id}$. The overview explains the same mechanism in Section 2.3, and the precise statement is Theorem 7.6,

which is proved in Part IV. The reduction of band matrices to a one-dimensional supersymmetric sigma model, with the scaling parameter N/W^2, is due to Fyodorov and Mirlin [53]; see also [94, (5.39)–(5.41)], [95, Sections 2.5.4 and 4.3], and [39].

By Proposition 4.11, the ratio $T_N^{(k)}$ is a ratio of integrals over chains $S_1, \dots, S_N$ of $(2k|2k)$ supermatrices. The source-free onsite weight is stationary where $S = (E - S)^{-1}$. Writing $S = \kappa_E - \mathrm{i}s_E Q$, this equation becomes $s_E^2(\mathrm{Id} - Q^2) = 0$, so that the critical points are the supermatrices $\kappa_E - \mathrm{i}s_E Q$ with $Q^2 = \mathrm{Id}$, as in (4.24). The chain stays near the balanced component of the saddle manifold of Section 4.6, on which Q lies in the orbit of Λ under conjugation, and its tangent variables move along this component. The target space $\mathcal{M}_k$ is this orbit, with its odd directions, and the balanced component of Section 4.6 is the set of supermatrices $\kappa_E - \mathrm{i}s_E Q$ with $Q \in \mathcal{M}_k$. To first order, each link moves the tangent variables by a heat semigroup at time $(4s_E^2 W^2)^{-1}$, since the link factor $e^{-W^2 \operatorname{STr}(S_{i+1} - S_i)^2/2}$ is a Gaussian in the distance between neighbors, the heat kernel $k_{\tau_{\mathrm{phys}}}$ of Definition 4.9. The N links therefore run this semigroup for the time $X_N(E)$ of (2.3). Since $Ns_E/a_N = \pi$, the source factors of the N sites combine to $\exp(-\pi \int_0^1 \operatorname{STr}(Q(t)\Gamma)\,\mathrm{d}t)$, where $Q(t)$ is the tangent variable at the site tN.

The odd directions of $\mathcal{M}_k$ matter for the following reason. By Remark 4.12, the onsite weight and the links are invariant under conjugation by all even invertible supermatrices, including those with odd entries. On $\mathcal{M}_k$, the infinitesimal odd symmetries give the localization identity Lemma 5.19, by which the integral of a superfunction annihilated by an odd symmetry is its value at the base point Λ. This identity implies the normalization $\mathcal{Z}_{X,\varepsilon}^{(k)}(0, \varepsilon\mathbf{1}) = 1$, which matches $T_N^{(k)}(0, \varepsilon\mathbf{1}) = 1$, and at rank one it produces the corner terms from which the kernels of Definition 1.4 arise (Section 6).

A sigma model is a field theory whose field takes values in a symmetric space, here $\mathcal{M}_k$, and the wire is the interval $[0,1]$ of scaled positions $t = i/N$ along the chain. Formally, the partition function of Definition 5.16 is a path integral over fields $Q\colon [0,1] \to \mathcal{M}_k$, whose transfer operator over a length t of the wire is $e^{-t\mathsf{G}_{X,\Gamma}}$, with the generator $\mathsf{G}_{X,\Gamma} = X\mathsf{C}_k + \pi \operatorname{STr}(Q\Gamma)$ of (5.23). This picture omits two difficulties, which occupy most of Part IV. The tangent and normal variables of the chain are coupled, and the superfunctions involved have no positivity, so that a Feynman–Kac formula on $\mathcal{M}_k$ is not available. Further, the chain variables lie on a contour rather than on the saddle manifold, and the weights are complex away from it.

5.2 The target space

Throughout this section, $k \geq 1$ is fixed. We first describe the superspace $\mathcal{M}_k = U(k,k|2k)/(U(k|k) \times U(k|k))$ through its underlying manifold and its odd coordinates, and we then define the field Q, the symmetries, the invariant Berezin integral, and the Casimir operator. We give concrete definitions in global coordinates, and we cite the

literature for the facts from the theory of homogeneous supermanifolds that we use. Physical background on these spaces can be found in [39] and [139].

Supermatrices

We use the supermatrices of Definition 3.10 of size $(2k|2k)$. We index their rows and columns by pairs (a,σ), where $a \in [2k]$ is a *flavor* and $\sigma \in \{\mathrm{B},\mathrm{F}\}$ is a *grading*. As in Section 4, the flavors $a \le k$ are retarded and the flavors $a > k$ advanced, and we write $|(a,\sigma)| = 0$ if $\sigma = \mathrm{B}$ and $|(a,\sigma)| = 1$ if $\sigma = \mathrm{F}$. For a supermatrix M, we let M^{BB}, M^{BF}, M^{FB}, and M^{FF} denote its four blocks of size $2k \times 2k$, and we let M_{RR}, M_{RA}, M_{AR}, and M_{AA} denote its four flavor blocks, which are supermatrices of size $(k|k)$. For instance, $M^{\mathrm{FF}}_{\mathrm{RA}}$ is the $k \times k$ matrix of the entries with retarded row flavor, advanced column flavor, and grading F. Recall the supermatrix $\Lambda = \mathrm{diag}(\Lambda_{\mathrm{B}}|\Lambda_{\mathrm{F}})$ of Section 4, with $\Lambda_{\mathrm{B}} = \Lambda_{\mathrm{F}} = \mathrm{diag}(\mathrm{Id}_k, -\mathrm{Id}_k)$. We let Id_{R} and Id_{A} denote the diagonal supermatrices with entry one at the retarded, respectively advanced, flavors in each grading, and zero elsewhere. Then $\Lambda = \mathrm{Id}_{\mathrm{R}} - \mathrm{Id}_{\mathrm{A}}$.

We let $\mathfrak{g}$ denote the space of all $(2k|2k)$ supermatrices with complex entries. We let $\mathfrak{g}_0$ denote the subspace of supermatrices with vanishing blocks M^{BF} and M^{FB}, and $\mathfrak{g}_1$ the subspace with vanishing blocks M^{BB} and M^{FF}, and we call the elements of $\mathfrak{g}_0$ even and those of $\mathfrak{g}_1$ odd. For homogeneous $Y, Y' \in \mathfrak{g}$ of parities $|Y|$ and $|Y'|$, we define the *bracket* $[Y,Y'] = YY' - (-1)^{|Y||Y'|}Y'Y$, and we extend it bilinearly. We call $\mathfrak{g}$, with this bracket, the Lie superalgebra $\mathfrak{gl}(2k|2k)$. All facts about it that we use are proved below. In this section and in Section 18.4, the letters Y and Y' denote elements of $\mathfrak{g}$; the local sources Y_i of Section 4 do not occur there. We set

$$\mathfrak{k} = \{Y \in \mathfrak{g} : \Lambda Y = Y\Lambda\}, \qquad \mathfrak{m} = \{Y \in \mathfrak{g} : \Lambda Y = -Y\Lambda\}, \tag{5.1}$$

and $\mathfrak{k}_j = \mathfrak{k} \cap \mathfrak{g}_j$, $\mathfrak{m}_j = \mathfrak{m} \cap \mathfrak{g}_j$ for $j \in \{0,1\}$. Thus $\mathfrak{k}$ consists of the supermatrices whose flavor blocks Y_{RA} and Y_{AR} vanish, $\mathfrak{m}$ of those whose blocks Y_{RR} and Y_{AA} vanish, and $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}$. The space $\mathfrak{m}_1$, of dimension $4k^2$, consists of the odd supermatrices supported on the flavor blocks RA and AR. We equip $\mathfrak{m}_1$ with the Hilbert–Schmidt inner product $\langle Y, Y' \rangle = \mathrm{Tr}(Y^*Y')$. For a complex matrix M, we write $\|M\|_{\mathrm{op}}$ for its operator norm and $\|M\|_{\mathrm{HS}}$ for its Hilbert–Schmidt norm.

For example, at rank one we order the four indices as $(\mathrm{RB}, \mathrm{RF}, \mathrm{AB}, \mathrm{AF})$, as in Section 6. Then $\Lambda = \mathrm{diag}(1,1,-1,-1)$, and the elements of $\mathfrak{k}$ and $\mathfrak{m}$ are the matrices

$$\begin{pmatrix} a_1 & \alpha_1 & 0 & 0 \\ \beta_1 & b_1 & 0 & 0 \\ 0 & 0 & a_2 & \alpha_2 \\ 0 & 0 & \beta_2 & b_2 \end{pmatrix} \in \mathfrak{k}, \qquad \begin{pmatrix} 0 & 0 & a_3 & \alpha_3 \\ 0 & 0 & \beta_3 & b_3 \\ a_4 & \alpha_4 & 0 & 0 \\ \beta_4 & b_4 & 0 & 0 \end{pmatrix} \in \mathfrak{m},$$

with complex entries, where the a_j and b_j occupy the even positions and the α_j and β_j the odd ones. Thus $\mathfrak{m}_0$ and $\mathfrak{m}_1$ are both four-dimensional, and an element of $\mathfrak{m}_1$ couples

the retarded bosonic index with the advanced fermionic one, and the retarded fermionic index with the advanced bosonic one.

We let $U(k,k)$ denote the group of complex $2k \times 2k$ matrices g with $g^* \Lambda_{\mathrm{B}} g = \Lambda_{\mathrm{B}}$, and we set

$$G_0 = \big\{ \operatorname{diag}(g_{\mathrm{B}} | g_{\mathrm{F}}) : g_{\mathrm{B}} \in U(k,k),\ g_{\mathrm{F}} \in U(2k) \big\}, \qquad K_0 = \{ g \in G_0 : g\Lambda = \Lambda g \}.$$

The group K_0 consists of the elements of G_0 whose flavor blocks RA and AR vanish. It is isomorphic to $U(k)^4$, its elements are unitary, and conjugation by $\mathsf{k} \in K_0$ preserves $\mathfrak{m}_1$ and its inner product.

The underlying manifold

We define $\mathcal{M}_k^0$ to be the set of supermatrices $g\Lambda g^{-1}$ with $g \in G_0$, and we call it the *underlying manifold* of $\mathcal{M}_k$. The map $gK_0 \mapsto g\Lambda g^{-1}$ is a bijection from G_0/K_0 onto $\mathcal{M}_k^0$, through which $\mathcal{M}_k^0$ is a smooth manifold of dimension $4k^2$. A point $Q \in \mathcal{M}_k^0$ is block diagonal, $Q = \operatorname{diag}(Q^{\mathrm{BB}} | Q^{\mathrm{FF}})$, and $\mathcal{M}_k^0 = H_k \times S_k$, where

$$H_k = \{ g\Lambda_{\mathrm{B}} g^{-1} : g \in U(k,k) \} \cong U(k,k)/(U(k) \times U(k)),$$
$$S_k = \{ u\Lambda_{\mathrm{F}} u^{-1} : u \in U(2k) \} \cong U(2k)/(U(k) \times U(k)).$$

We equip $\mathcal{M}_k^0$ with the Riemannian metric whose square norm of a tangent vector $\dot{Q}$ at Q is

$$-\tfrac{1}{2} \operatorname{STr}(\dot{Q}^2) = -\tfrac{1}{2} \operatorname{Tr}\big((\dot{Q}^{\mathrm{BB}})^2\big) + \tfrac{1}{2} \operatorname{Tr}\big((\dot{Q}^{\mathrm{FF}})^2\big), \tag{5.2}$$

and we let dvol denote its volume measure and $\Delta_{\mathcal{M}_k^0}$ its Laplace–Beltrami operator, which is nonpositive. The metric is invariant under the action $Q \mapsto gQg^{-1}$ of G_0. The space H_k is the Riemannian symmetric space of the noncompact group $U(k,k)$, and S_k is the Grassmannian of k-dimensional subspaces of $\mathbb{C}^{2k}$. To see that (5.2) is positive definite, it suffices by invariance to consider $Q = \Lambda$. There the tangent vectors have the form $\dot{Q} = [Y, \Lambda] = -2\Lambda Y$, where $Y \in \mathfrak{m}_0$ satisfies $Y_{\mathrm{AR}}^{\mathrm{BB}} = (Y_{\mathrm{RA}}^{\mathrm{BB}})^*$ and $Y_{\mathrm{AR}}^{\mathrm{FF}} = -(Y_{\mathrm{RA}}^{\mathrm{FF}})^*$, and for these

$$-\tfrac{1}{2} \operatorname{STr}(\dot{Q}^2) = 4 \operatorname{Tr}\big(Y_{\mathrm{RA}}^{\mathrm{BB}} (Y_{\mathrm{RA}}^{\mathrm{BB}})^*\big) + 4 \operatorname{Tr}\big(Y_{\mathrm{RA}}^{\mathrm{FF}} (Y_{\mathrm{RA}}^{\mathrm{FF}})^*\big).$$

At rank one, H_1 is the set of matrices $\left(\begin{smallmatrix} B & \sqrt{B^2-1}\, e^{\mathrm{i}\theta_{\mathrm{B}}} \\ -\sqrt{B^2-1}\, e^{-\mathrm{i}\theta_{\mathrm{B}}} & -B \end{smallmatrix}\right)$ with $B \geq 1$, and S_1 is the set of matrices $\left(\begin{smallmatrix} F & \mathrm{i}\sqrt{1-F^2}\, e^{\mathrm{i}\theta_{\mathrm{F}}} \\ -\mathrm{i}\sqrt{1-F^2}\, e^{-\mathrm{i}\theta_{\mathrm{F}}} & -F \end{smallmatrix}\right)$ with $F \in [-1,1]$. In the coordinates $B = \cosh r$ and $F = \cos\vartheta$, the metric (5.2) takes the form $\mathrm{d}r^2 + \sinh^2 r\, \mathrm{d}\theta_{\mathrm{B}}^2$ on H_1 and $\mathrm{d}\vartheta^2 + \sin^2\vartheta\, \mathrm{d}\theta_{\mathrm{F}}^2$ on S_1. Thus H_1 is the hyperbolic plane of curvature -1 in the hyperboloid model, S_1 is the unit sphere, and the coordinates B and F are those of Definition 1.4. In these coordinates, $\operatorname{dvol} = \mathrm{d}B\, \mathrm{d}\theta_{\mathrm{B}}\, \mathrm{d}F\, \mathrm{d}\theta_{\mathrm{F}}$, and on functions of (B,F) the operator $\Delta_{\mathcal{M}_1^0}$ acts as Δ_D.

We define the *weight* v on $\mathcal{M}_k^0$ by

$$v(Q) = 1 + \tfrac{1}{2}\operatorname{Tr}(\Lambda_{\mathrm{B}} Q^{\mathrm{BB}}). \tag{5.3}$$

At rank one, $v = 1 + B$. The following lemma collects the properties of v that we use.

Lemma 5.1. *(a) For all $Q \in \mathcal{M}_k^0$, we have $\frac{1}{2}\operatorname{Tr}(\Lambda_{\mathrm{B}} Q^{\mathrm{BB}}) \geq k$, with equality if and only if $Q^{\mathrm{BB}} = \Lambda_{\mathrm{B}}$, and $\frac{1}{2}\operatorname{Tr}(\Lambda_{\mathrm{F}} Q^{\mathrm{FF}}) \leq k$, with equality if and only if $Q^{\mathrm{FF}} = \Lambda_{\mathrm{F}}$. Moreover, $\|Q^{\mathrm{BB}}\|_{\mathrm{HS}} \leq 2v(Q)$ and $\|Q^{\mathrm{FF}}\|_{\mathrm{op}} = 1$.*

(b) For all integers $m \geq 1$, there exists C_m such that $|\nabla^m v| \leq C_m v$ on $\mathcal{M}_k^0$, where ∇ is the Levi-Civita connection.

(c) There exist C and q_1 such that $\operatorname{vol}(\{v \leq R\}) \leq CR^{q_1}$ for all $R \geq 1$.

Proof. We begin with (a). Let $Q^{\mathrm{BB}} = g\Lambda_{\mathrm{B}} g^{-1}$ with $g \in U(k,k)$, and write g in flavor blocks a, b, c, e. Since $g^{-1} = \Lambda_{\mathrm{B}} g^* \Lambda_{\mathrm{B}}$, we have

$$\Lambda_{\mathrm{B}} Q^{\mathrm{BB}} = \Lambda_{\mathrm{B}} g g^* \Lambda_{\mathrm{B}} \geq 0, \qquad \operatorname{Tr}(\Lambda_{\mathrm{B}} Q^{\mathrm{BB}}) = \operatorname{Tr}(gg^*).$$

The relation $g^* \Lambda_{\mathrm{B}} g = \Lambda_{\mathrm{B}}$ implies that $a^*a - c^*c = \mathrm{Id}_k$ and $e^*e - b^*b = \mathrm{Id}_k$, and therefore

$$\operatorname{Tr}(gg^*) = 2k + 2\operatorname{Tr}(b^*b) + 2\operatorname{Tr}(c^*c) \geq 2k.$$

Equality holds if and only if $b = c = 0$, or equivalently if g commutes with Λ_{B}. Since $\Lambda_{\mathrm{B}} Q^{\mathrm{BB}}$ is positive semidefinite, we have

$$\|Q^{\mathrm{BB}}\|_{\mathrm{HS}} = \|\Lambda_{\mathrm{B}} Q^{\mathrm{BB}}\|_{\mathrm{HS}} \leq \operatorname{Tr}(\Lambda_{\mathrm{B}} Q^{\mathrm{BB}}) = 2(v-1).$$

The matrix $Q^{\mathrm{FF}} = u\Lambda_{\mathrm{F}} u^*$ is Hermitian and unitary, and $\operatorname{Tr}(\Lambda_{\mathrm{F}} Q^{\mathrm{FF}}) \leq 2k$, with equality if and only if the unitary matrix $\Lambda_{\mathrm{F}} u \Lambda_{\mathrm{F}} u^*$ is the identity.

We now prove (b). Fix $g' = \operatorname{diag}(g'_{\mathrm{B}} | g'_{\mathrm{F}}) \in G_0$, and consider the isometry $\iota_{g'}(Q) = g' Q g'^{-1}$. Then

$$v \circ \iota_{g'}(Q) = 1 + \tfrac{1}{2}\operatorname{Tr}(A_{g'} Q^{\mathrm{BB}}), \qquad A_{g'} = (g'_{\mathrm{B}})^{-1} \Lambda_{\mathrm{B}} g'_{\mathrm{B}} = \Lambda_{\mathrm{B}} (g'_{\mathrm{B}})^* g'_{\mathrm{B}}.$$

The function $v \circ \iota_{g'}$ is the restriction to $\mathcal{M}_k^0$ of an affine function of the entries of Q. Its covariant derivatives of order m at Λ are therefore bounded by

$$C_m \|A_{g'}\|_{\mathrm{HS}} \leq C_m \operatorname{Tr}((g'_{\mathrm{B}})^* g'_{\mathrm{B}}) = 2C_m \big(v(g' \Lambda g'^{-1}) - 1\big),$$

with C_m independent of g'. Since $\iota_{g'}$ is an isometry mapping Λ to $g'\Lambda g'^{-1}$, and G_0 acts transitively, we obtain (b).

Finally, we prove (c). It suffices to consider H_k, since S_k is compact. Every point of H_k has the form $Q^{\mathrm{BB}} = e^Y \Lambda_{\mathrm{B}} e^{-Y}$ with $Y = \left(\begin{smallmatrix} 0 & w \\ w^* & 0 \end{smallmatrix}\right)$, since H_k is a symmetric space of

noncompact type and its exponential map at Λ_{B} is onto. Let $\sigma_1, \dots, \sigma_k$ be the singular values of w.

The curve $s \mapsto e^{sY}\Lambda_{\mathrm{B}}e^{-sY}$, $0 \le s \le 1$, is a geodesic, whose length r is the distance from Λ_{B} to Q^{BB}. Using

$$\Lambda_{\mathrm{B}}Q^{\mathrm{BB}} = \Lambda_{\mathrm{B}}e^{2Y}\Lambda_{\mathrm{B}} = e^{-2Y},$$

we find

$$r = 2\Big(\sum_{j=1}^{k}\sigma_j^2\Big)^{1/2}, \qquad v = 1 + \sum_{j=1}^{k}\cosh(2\sigma_j) \ge \cosh(r/\sqrt{k}).$$

Thus $\{v \le R\}$ is contained in the product of S_k with the ball of radius $\sqrt{k}\log(2R)$ about Λ_{B} in H_k. Since H_k is homogeneous, its sectional curvature is bounded below, and the Bishop–Gromov inequality bounds the volume of a ball of radius r by Ce^{Cr}. Together with the inclusion above, this implies (c), since $e^{C\sqrt{k}\log(2R)} = (2R)^{C\sqrt{k}}$. □

Superfunctions and the field Q

Let $\bigwedge_{4k^2}$ denote the Grassmann algebra with odd generators $\chi_1, \dots, \chi_{4k^2}$ (Definition 3.1). We fix an orthonormal basis $e_1, \dots, e_{4k^2}$ of $\mathfrak{m}_1$, and we define the *generic odd tangent supermatrix* to be $\Theta = \sum_a \chi_a e_a$. Its nonzero entries lie in the blocks BF and FB and are linear combinations of the χ_a. In particular, Θ is an even supermatrix in the sense of Definition 3.10. For $\mathsf{k} \in K_0$, we let α_{k} denote the automorphism of the algebra $\bigwedge_{4k^2}$ with $\sum_a \alpha_{\mathsf{k}}(\chi_a)e_a = \mathsf{k}\Theta\mathsf{k}^{-1}$, which replaces Θ by $\mathsf{k}\Theta\mathsf{k}^{-1}$.

Definition 5.2. A *superfunction* on $\mathcal{M}_k$ is a smooth map $f\colon G_0 \to \bigwedge_{4k^2}$ such that $f(g\mathsf{k}) = \alpha_{\mathsf{k}}(f(g))$ for all $g \in G_0$ and $\mathsf{k} \in K_0$. We write $f(g;\Theta)$ for $f(g)$ to indicate the dependence on the odd variables. The *field* Q is the matrix of superfunctions

$$Q(g;\Theta) = g\,e^{\Theta}\Lambda e^{-\Theta}g^{-1} = g\Lambda e^{-2\Theta}g^{-1}. \tag{5.4}$$

The second form of (5.4) uses $\Theta\Lambda = -\Lambda\Theta$. Since $\mathsf{k}\Lambda\mathsf{k}^{-1} = \Lambda$, we have

$$Q(g\mathsf{k};\Theta) = g\Lambda e^{-2\mathsf{k}\Theta\mathsf{k}^{-1}}g^{-1} = \alpha_{\mathsf{k}}\big(Q(g;\Theta)\big),$$

and Q is therefore a matrix of superfunctions. Its entries have the parities of their positions, so Q is an even supermatrix. It satisfies

$$Q^2 = \mathrm{Id}, \qquad \mathrm{STr}\,Q = 0. \tag{5.5}$$

For the second identity, note that $\mathrm{STr}\,Q = \mathrm{STr}(\Lambda e^{-2\Theta})$ and that $\mathrm{STr}(\Lambda\Theta^m) = 0$ for all $m \ge 1$. Indeed, for odd m the supermatrix $\Lambda\Theta^m$ has vanishing diagonal blocks, and for even $m \ge 2$ the cyclicity of the supertrace (Lemma 3.11(c)) and $\Theta\Lambda = -\Lambda\Theta$ give $\mathrm{STr}(\Lambda\Theta^m) = \mathrm{STr}(\Theta\Lambda\Theta^{m-1}) = -\mathrm{STr}(\Lambda\Theta^m)$.

Definition 5.2 is the global description of $\mathcal{M}_k$ that we use, and we call (g, Θ) the *regular coordinates*. We call $g\Lambda g^{-1} \in \mathcal{M}_k^0$ the body point of $(g;\Theta)$. The body $f_\emptyset$ of a superfunction is then a function on $\mathcal{M}_k^0$, and $Q_\emptyset(g) = g\Lambda g^{-1}$. Let $\mathcal{O} \subset \mathcal{M}_k^0$ be open. We call a smooth map $\mathsf{s}\colon \mathcal{O} \to G_0$ with $\mathsf{s}(x)\Lambda\mathsf{s}(x)^{-1} = x$ for all $x \in \mathcal{O}$ a *section* over $\mathcal{O}$. For a section s, the map $f \mapsto f \circ \mathsf{s}$ identifies the superfunctions on $\mathcal{M}_k$, restricted to $\mathcal{O}$, with the superfunctions on the superdomain $\mathcal{O}^{4k^2|4k^2}$ of Definition 3.2, with odd coordinates χ_a. A second section has the form sk for a smooth map $\mathsf{k}\colon \mathcal{O} \to K_0$, and the two identifications differ by the unitary change of odd coordinates $\Theta \mapsto \mathsf{k}\Theta\mathsf{k}^{-1}$.

The regular coordinates identify the superfunctions with the sections of a vector bundle over $\mathcal{M}_k^0$ with fiber $\bigwedge_{4k^2}$, and all analytic estimates below are based on this structure. The word regular distinguishes them from the polar coordinates of Section 6, which are singular at a point. The regular coordinates are nonsingular everywhere, and integrals in them require only integration by parts, without boundary terms (compare Section 3.6).

We use the conventions of Section 3 for these superfunctions. Coefficients are written to the left of the monomials in the χ_a, and auxiliary odd generators, such as the parameters ϑ_Y of Section 5.2, are treated as coefficients (Section 3.2). If $a_1, \dots, a_m$ are even superfunctions and ϕ is a smooth function on a neighborhood of the range of their bodies, then the superfunction $\phi(a_1, \dots, a_m)$ is defined pointwise by Definition 3.3; this defines, for instance, e^{-cd_k} and $\psi(d_1/r)$ below, and compositions of such functions are computed by Lemma 3.4(d). Since the number $4k^2$ of the generators χ_a is even, the Berezin integral in the χ_a commutes with Berezin integrals in other generators (Lemma 3.6(f)).

Remark 5.3. Every superfunction on $\mathcal{M}_k$ is, near each point of $\mathcal{M}_k^0$, a smooth function of the entries of Q, in the sense of Definition 3.3. Indeed, near a point of $\mathcal{M}_k^0$, the bodies of suitable even entries of Q form a coordinate system of $\mathcal{M}_k^0$, since $\mathcal{M}_k^0$ is a submanifold of a space of matrices. By (5.4), the terms of degree one of the odd entries of Q form the matrix $-2g\Lambda\Theta g^{-1}$, which determines Θ, and their other terms have degree at least three in the χ_a. Solving for Θ degree by degree, we write the χ_a as polynomials in the odd entries of Q whose coefficients are smooth functions of the body point. The souls of the even entries are polynomials in the χ_a, so the bodies of the even entries are also smooth functions of the entries of Q, and inserting these expressions into $f = \sum_I f_I \chi^I$ proves the claim. For intuition, a superfunction on $\mathcal{M}_k$ may therefore be thought of as a function of a supermatrix Q with $Q^2 = \mathrm{Id}$, as in the physics literature, and a derivation of the algebra of superfunctions is determined by its values on the entries of Q.

Symmetries

We use the name $U(k,k|2k)$ only for the pair consisting of the group G_0 and the Lie superalgebra $\mathfrak{g}$. The group G_0 acts on $\mathcal{M}_k^0$ by $Q \mapsto g'Qg'^{-1}$, and on superfunctions by left translation of g in the regular coordinates. In the odd directions there are no

group elements with complex entries, and the symmetries are described infinitesimally, by derivations of the algebra of superfunctions. For each $Y \in \mathfrak{g}$, we now construct the derivation $\widehat{Y}$ that moves Q by its commutator with Y. For even Y, it is the derivative of the action of e^{sY}, and for odd Y, it is the odd analogue, written with an odd parameter. In the language of supermanifolds, the $\widehat{Y}$ are the fundamental vector fields of the action of the supergroup on the homogeneous superspace $\mathcal{M}_k$; for textbook accounts of supergroups and their actions, see [8, 132]. We construct them directly, in the regular coordinates, and prove the properties that we use. Let $Y \in \mathfrak{g}$ be homogeneous. We let ϑ_Y be a real parameter if Y is even, and we work to first order in it. We let ϑ_Y be an auxiliary odd generator if Y is odd, in the sense of Section 3.3. Then $\vartheta_Y Y$ is an even supermatrix and $\vartheta_Y^2 = 0$.

Lemma 5.4. *(a) For every homogeneous $Y \in \mathfrak{g}$, there exists a unique derivation $\widehat{Y}$ of the algebra of superfunctions on $\mathcal{M}_k$, of the parity of Y, such that*

$$\vartheta_Y \widehat{Y} Q = [\vartheta_Y Y, Q]. \tag{5.6}$$

(b) For $g' \in G_0$, let $(L_{g'} f)(g;\Theta) = f(g'^{-1}g;\Theta)$. Then $L_{g'} \widehat{Y} L_{g'}^{-1} = \widehat{g'Yg'^{-1}}$. If Y is even, then $\widehat{Y}Q$ is the matrix $[Y,Q]$, and if moreover Y lies in the Lie algebra of G_0, then $(\widehat{Y}f)(g;\Theta) = \frac{\mathrm{d}}{\mathrm{d}s} f(e^{sY}g;\Theta)$ at $s = 0$.

(c) Let $g \in G_0$ and $s \geq 0$, and let f be a superfunction whose coefficients of degree less than s in the χ_a vanish on a neighborhood of g. If $Y = \sum_a y_a e_a \in \mathfrak{m}_1$, then the terms of degree at most s of

$$(\widehat{gYg^{-1}} f)(g;\Theta) - \sum_a y_a \frac{\partial f}{\partial \chi_a}(g;\Theta)$$

vanish. If $Y \in \mathfrak{k}_1$, then the terms of degree at most s of $(\widehat{gYg^{-1}} f)(g;\Theta)$ vanish.

(d) For homogeneous $Y, Y' \in \mathfrak{g}$, we have

$$[\widehat{Y}, \widehat{Y'}] = -(-1)^{|Y||Y'|} \widehat{[Y,Y']},$$

where $[D,D'] = DD' - (-1)^{|D||D'|} D'D$ for operators D and D' of parities $|D|$ and $|D'|$.

Here and below, $\widehat{Y}Q$ denotes the matrix obtained by applying $\widehat{Y}$ to each entry of Q. A derivation of parity one is an odd derivation, which means that $\widehat{Y}(ff') = (\widehat{Y}f)f' + (-1)^{|f|} f \widehat{Y} f'$ for homogeneous f of parity $|f|$. By the uniqueness in (a), $\widehat{Y}$ depends linearly on Y, and we extend the definition to all $Y \in \mathfrak{g}$ by linearity. In words, (c) says that, at a given point, the odd fields $\widehat{gYg^{-1}}$ with $Y \in \mathfrak{m}_1$ act to leading order as the derivatives in the odd coordinates, while those with $Y \in \mathfrak{k}_1$ raise the degree.

Proof. We first construct $\widehat{Y}$. Fix $(g;\Theta)$, write $\vartheta = \vartheta_Y$, and adjoin ϑ to the coefficient algebra, working modulo ϑ^2 when it is even. Set $Z = e^{-\Theta}(\vartheta g^{-1} Y g) e^{\Theta}$, which is an even

supermatrix. We look for $A \in \mathfrak{m}_0$, $\Theta_1 \in \mathfrak{m}_1$, and $C \in \mathfrak{k}$, with coefficients in the original algebra $\bigwedge_{4k^2}$, such that

$$e^{-\Theta}(\vartheta A)e^{\Theta} + \frac{1 - e^{-\mathrm{ad}_\Theta}}{\mathrm{ad}_\Theta}(\vartheta\Theta_1) + \vartheta C = Z, \tag{5.7}$$

where $\mathrm{ad}_\Theta(Y') = \Theta Y' - Y'\Theta$, and the fraction is defined by its power series, which terminates since Θ is nilpotent. We solve for the variations $U = \vartheta A$, $V = \vartheta\Theta_1$, and $K = \vartheta C$, all of which are even supermatrices. On the decomposition $\mathfrak{g} = \mathfrak{m}_0 \oplus \mathfrak{m}_1 \oplus \mathfrak{k}$ with coefficients in the extended algebra, the left side is the identity plus a linear map that raises the degree in the original generators χ_a. Its inverse is a finite Neumann series. It preserves the ideal generated by ϑ, and gives a unique solution (U, V, K) linear in ϑ. Extracting ϑ on the left defines A, Θ_1, and C uniquely in the original algebra. Their coefficients in the χ_a are linear functions of $g^{-1}Yg$, with operator norms bounded in terms of k and the fixed basis. The parameter stays inside the matrix products in (5.7); for odd Y, this accounts for $e^{-\Theta}\vartheta = \vartheta e^{\Theta}$.

By the formula for the differential of the exponential map, (5.7) says that

$$e^{\vartheta Y} g e^{\Theta} = g(\mathrm{Id} + \vartheta A)e^{\Theta + \vartheta\Theta_1}(\mathrm{Id} + \vartheta C)$$

to first order in ϑ. We define $\widehat{Y}f$ by

$$f\big(g(\mathrm{Id} + \vartheta A); \Theta + \vartheta\Theta_1\big) = f(g;\Theta) + \vartheta(\widehat{Y}f)(g;\Theta), \tag{5.8}$$

where the left side is defined by Taylor expansion in the direction of the left-invariant vector field A on G_0 (Definition 3.3). Here A has coefficients in $\mathfrak{m}_0 \otimes \bigwedge_{4k^2}$, and $\mathfrak{m}_0$ is not contained in the Lie algebra of G_0. Since $\mathfrak{g}_0$ is the complexification of the Lie algebra of G_0, the left-invariant derivative $\frac{\mathrm{d}}{\mathrm{d}s} f(g e^{sA})$, defined for A in this Lie algebra, extends complex-linearly to $A \in \mathfrak{g}_0$, and the infinitesimal form of the equivariance under K_0 extends in the same way to $\mathfrak{k}_0$; this is the meaning of the Taylor expansion here and of the equivariance used below. The first-order substitution is multiplicative, so extracting ϑ on the left gives the graded Leibniz rule for $\widehat{Y}$, a derivation of the parity of Y. It maps superfunctions to superfunctions, since replacing (g, Θ) by $(g\mathsf{k}, \mathsf{k}^{-1}\Theta\mathsf{k})$ conjugates the variations (U, V, K) by k^{-1}, by uniqueness of the decomposition, and hence conjugates (A, Θ_1, C) in the same way. Since $\mathrm{Id} + \vartheta C$ commutes with Λ to first order, we have

$$Q(g(\mathrm{Id} + \vartheta A); \Theta + \vartheta\Theta_1) = e^{\vartheta Y} Q(g;\Theta) e^{-\vartheta Y},$$

which is (5.6).

For even Y, let $A_0 + C_0$ be the decomposition of $g^{-1}Yg \in \mathfrak{g}_0$ with $A_0 \in \mathfrak{m}_0$ and $C_0 \in \mathfrak{k}_0$. To first order in s, we have $(\mathrm{Id} + sC_0)e^{\Theta} = e^{\Theta + s[C_0,\Theta]}(\mathrm{Id} + sC_0)$. The solution of (5.7) is therefore $A = A_0$, $\Theta_1 = [C_0, \Theta]$, and $C = C_0$, and the equivariance of f under K_0 implies

that $\widehat{Y}f(g;\Theta) = \frac{\mathrm{d}}{\mathrm{d}s} f(e^{sY}g;\Theta)$ at $s=0$. The identity $L_{g'}\widehat{Y}L_{g'}^{-1} = \widehat{g'Yg'^{-1}}$ follows from $e^{\vartheta_Y Y} g'^{-1} g = g'^{-1} e^{\vartheta_Y g' Y g'^{-1}} g$.

It remains to prove uniqueness. It suffices to show that a derivation $\mathcal{D}$ with $\mathcal{D}Q = 0$ vanishes. By Remark 5.3, every superfunction f is locally a smooth function of the entries of Q, and the chain rule of Lemma 3.4(b) implies that $\mathcal{D}f = 0$.

We now prove (c). By (b), we have $(\widehat{gYg^{-1}}f)(g;\Theta) = (\widehat{Y}L_g^{-1}f)(\mathrm{Id};\Theta)$, and $L_g^{-1}f = f(g\,\cdot\,;\cdot)$ has the same vanishing coefficients near Id. It suffices to consider $g = \mathrm{Id}$. We solve (5.7) degree by degree in the χ_a, retaining ϑ in every matrix product. Since ϑA and $\vartheta\Theta_1$ are even and ϑ is odd, the coefficients of A are odd and those of Θ_1 are even. The equation in degree zero is $\vartheta(\Theta_1^{(0)} + C^{(0)}) = \vartheta Y$, hence $\Theta_1^{(0)} + C^{(0)} = Y$, where the superscript denotes the part of degree zero in the original generators. For $Y \in \mathfrak{m}_1$, we obtain $\Theta_1 = Y + O(\chi^2)$, and for $Y \in \mathfrak{k}_1$, we obtain $\Theta_1 = O(\chi^2)$, where $O(\chi^j)$ denotes terms of degree at least j. In both cases $A = O(\chi)$.

By (5.8), $(\widehat{Y}f)(\mathrm{Id};\Theta)$ is the sum of the left-invariant derivative of f in the direction A and of the derivative $\sum_a (\Theta_1)_a \partial f/\partial\chi_a$ in the odd direction $\Theta_1 = \sum_a (\Theta_1)_a e_a$. The first term, and the part $O(\chi^2)$ of Θ_1, contribute only terms of degree at least $s+1$, and the part Y of Θ_1 contributes $\sum_a y_a \partial f/\partial\chi_a$. This proves (c). The supplementary script `polar/check_regular_action.py` verifies the factorization, its action on Q, the changes of section, and these degree bounds at rank one for all sixteen matrix units.

Finally, we prove (d). Let $\vartheta = \vartheta_Y$ and $\vartheta' = \vartheta_{Y'}$ be independent parameters, and set $D = \vartheta\widehat{Y}$ and $D' = \vartheta'\widehat{Y}'$. These are even derivations with $DQ = [\vartheta Y, Q]$ and $D'Q = [\vartheta' Y', Q]$, and they commute with multiplication by constants. We obtain

$$(DD' - D'D)Q = [\vartheta' Y', [\vartheta Y, Q]] - [\vartheta Y, [\vartheta' Y', Q]] = -\big[[\vartheta Y, \vartheta' Y'], Q\big] = -\big[\vartheta\vartheta'[Y, Y'], Q\big],$$

where we used the Jacobi identity for commutators of the even supermatrices ϑY, $\vartheta' Y'$, and Q, and the identity

$$\vartheta Y \vartheta' Y' - \vartheta' Y' \vartheta Y = \vartheta\vartheta'\big(YY' - (-1)^{|Y||Y'|} Y'Y\big).$$

The derivation $DD' - D'D + \vartheta\vartheta'\widehat{[Y, Y']}$ therefore vanishes on Q, and the uniqueness in (a) shows that

$$DD' - D'D = -\vartheta\vartheta'\,\widehat{[Y, Y']}.$$

On the other hand, moving the parameters to the left, we find

$$DD' - D'D = (-1)^{|Y||Y'|}\vartheta\vartheta'\,[\widehat{Y}, \widehat{Y}'].$$

Comparing the coefficients of $\vartheta\vartheta'$ proves (d). □

We call a superfunction f *invariant* under a subspace $\mathfrak{h} \subset \mathfrak{g}$ if $\widehat{Y}f = 0$ for all homogeneous $Y \in \mathfrak{h}$. The function

$$d_k(Q) = \tfrac{1}{2}\,\mathrm{STr}(\Lambda Q) \tag{5.9}$$

is invariant under $\mathfrak{k}$. Indeed, for $Y \in \mathfrak{k}$, the cyclicity of the supertrace gives

$$\vartheta_Y \widehat{Y} d_k = \tfrac{1}{2} \operatorname{STr}(\Lambda[\vartheta_Y Y, Q]) = \tfrac{1}{2} \operatorname{STr}([\Lambda, \vartheta_Y Y] Q) = 0.$$

We have $(d_k)_\varnothing = \frac{1}{2} \operatorname{Tr}(\Lambda_{\mathrm{B}} Q^{\mathrm{BB}}) - \frac{1}{2} \operatorname{Tr}(\Lambda_{\mathrm{F}} Q^{\mathrm{FF}})$ on $\mathcal{M}_k^0$. By Lemma 5.1(a), it is nonnegative and vanishes only at $Q = \Lambda$, and

$$v - 1 - k \le (d_k)_\varnothing \le v - 1 + k. \tag{5.10}$$

The following lemma shows that a superfunction that is invariant under $\mathfrak{g}$ is constant. We use it in the uniqueness of the invariant integral below.

Lemma 5.5. *Let f be a superfunction on $\mathcal{M}_k$ that is invariant under $\mathfrak{g}$. Then f is constant. This also holds for a superfunction f over a connected open set $\mathcal{O} \subset \mathcal{M}_k^0$, that is, for a smooth map from the preimage of $\mathcal{O}$ in G_0 to $\bigwedge_{4k^2}$ with the equivariance of Definition 5.2.*

Proof. Let $\tilde{\mathcal{O}} \subset G_0$ be the preimage of $\mathcal{O}$, where $\mathcal{O} = \mathcal{M}_k^0$ in the first statement. It is connected, since $\mathcal{O}$ and K_0 are connected. For Y in the Lie algebra of G_0, Lemma 5.4(b) implies that $\frac{\mathrm{d}}{\mathrm{d}s} f(e^{sY} g; \Theta) = 0$ whenever $e^{sY} g \in \tilde{\mathcal{O}}$. The vectors Yg span the tangent space of G_0 at g, and therefore $f(\,\cdot\,;\Theta)$ is constant on $\tilde{\mathcal{O}}$. Let f_n denote the part of f of degree n in the χ_a. Suppose that f has a nonzero part of positive degree, and let $n \ge 1$ be the smallest such degree. The superfunction $f - f_0$ is invariant under $\mathfrak{g}$, since f_0 is constant, and its coefficients of degree less than n vanish.

Fix $g_0 \in \tilde{\mathcal{O}}$ and $Y = \sum_a y_a e_a \in \mathfrak{m}_1$. Since $\widehat{g_0 Y g_0^{-1}}(f - f_0) = 0$, Lemma 5.4(c), applied to $f - f_0$ at g_0 with $s = n$, shows that the terms of degree at most n of $\sum_a y_a\, \partial(f - f_0)/\partial\chi_a(g_0;\Theta)$ vanish. Its terms of degree $n-1$ give

$$\sum_a y_a \frac{\partial f_n}{\partial \chi_a}(g_0;\Theta) = 0 \qquad \text{for all } y \in \mathbb{C}^{4k^2}.$$

An element of $\bigwedge_{4k^2}$ of degree $n \ge 1$ whose derivatives in all the χ_a vanish is zero, and therefore $f_n(g_0;\cdot) = 0$. Since f_n does not depend on g, this contradicts the choice of n. □

The invariant integral and the Casimir operator

This subsection defines the invariant Berezin integral and the Casimir operator C_k of the sigma model. In Lemma 16.5, Part IV identifies C_k, up to a constant factor, with the heat generator of the chain restricted to the saddle manifold.

By Lemma 5.4, the action of G_0 only moves the body point in regular coordinates. A Berezin integral on $\mathcal{M}_k$ that is invariant under G_0 therefore has the form

$$\int f = \int_{\mathcal{M}_k^0} \Big(\int \rho(\Theta) f(g;\Theta)\, \mathrm{d}\chi \Big)\, \mathrm{dvol}(g \Lambda g^{-1}), \tag{5.11}$$

where $\int \cdot \, \mathrm{d}\chi$ is the Berezin integral of Definition 3.5 in the generators χ_a, and $\rho \in \bigwedge_{4k^2}$ is even and invariant under the automorphisms α_{k}. The inner integral is a function of $g\Lambda g^{-1}$, by Lemma 3.6(c), since α_{k} is induced by a linear map of $\mathfrak{m}_1$ of determinant one. (Conjugation by $\mathsf{k} = \mathrm{diag}(\mathsf{k}_{\mathrm{RB}}, \mathsf{k}_{\mathrm{AB}} | \mathsf{k}_{\mathrm{RF}}, \mathsf{k}_{\mathrm{AF}})$ acts on the four blocks of $\mathfrak{m}_1$ with determinants $\det(\mathsf{k}_{\mathrm{RB}})^k \det(\mathsf{k}_{\mathrm{AF}})^{-k}$, $\det(\mathsf{k}_{\mathrm{RF}})^k \det(\mathsf{k}_{\mathrm{AB}})^{-k}$, and their inverses.) As in Definition 3.5, a superfunction f is called integrable if $|f|_{1;1} \in L^1(\,\mathrm{dvol})$, where $|f|_{1;1}$ denotes the ℓ^1 norm of the coefficients of f, one of the coefficient norms of Section 5.2. For such f, the integral (5.11) converges absolutely.

Definition 5.6. We define the *invariant Berezin integral* $f \mapsto \int f$ to be the functional (5.11) with the even element ρ chosen such that $\int \widehat{Y} f = 0$ for all homogeneous $Y \in \mathfrak{g}$ and all compactly supported superfunctions f, and such that

$$\int e^{-d_k} = 1. \tag{5.12}$$

We define the *Berezin pairing* $\mathcal{B}_k(f, f') = \int ff'$ for superfunctions f and f' such that ff' is integrable.

There exists an even element ρ with nonzero constant body such that $\int \widehat{Y} f = 0$ for all homogeneous Y and all compactly supported f. In the terminology of [2], this says that the homogeneous superspace $\mathcal{M}_k$ is unimodular. It follows from [2, Corollary 4.12 and Theorem 4.13], which together show that such an element exists if and only if the Berezinian of $(\mathfrak{g}/\mathfrak{k})^*$ is a trivial module of the stabilizer; see also [139] for the Riemannian symmetric superspaces of random matrix theory. For the invariant integral, this means that the stabilizer of Λ, which acts on the Berezin densities on the tangent space $\mathfrak{g}/\mathfrak{k} \cong \mathfrak{m}$ (the elements of the Berezinian of $(\mathfrak{g}/\mathfrak{k})^*$) through the Berezinian of its action on $\mathfrak{m}$, leaves these densities unchanged, so that a density at Λ is transported consistently by the action of $U(k,k|2k)$ and defines an invariant integral of the form (5.11). Since the stabilizer of Λ is connected, this criterion reduces to the condition that, for every $Y \in \mathfrak{k}$, the linear map $Y' \mapsto [Y, Y']$ of $\mathfrak{m}$ has supertrace zero, where the supertrace of a parity-preserving linear map of $\mathfrak{m}$ is its trace on $\mathfrak{m}_0$ minus its trace on $\mathfrak{m}_1$. For a diagonal $Y \in \mathfrak{k}_0$ with diagonal entries y_i, this map multiplies the entry Y'_{ij} of $Y' \in \mathfrak{m}$ by $y_i - y_j$, and its supertrace is $\sum_{i,j} (-1)^{|i|+|j|}(y_i - y_j)$, summed over the pairs of indices in different retarded and advanced sectors. This sum vanishes, since each sector has k even and k odd indices. The off-diagonal entries of $Y \in \mathfrak{k}_0$ do not contribute to the trace, and for odd Y the map exchanges $\mathfrak{m}_0$ and $\mathfrak{m}_1$ and has supertrace zero.

The element ρ is unique up to a constant factor. Indeed, if ρ' is a second one, let $\sigma = \rho' \rho^{-1}$, regarded as a superfunction that is constant in g. For compactly supported f, the invariance of both integrals gives

$$\int (\widehat{Y}\sigma) f = \int \widehat{Y}(\sigma f) - \int \sigma \, \widehat{Y} f = 0,$$

where $\int \sigma \widehat{Y} f$ is the integral of $\widehat{Y} f$ with the density ρ'. Hence $\widehat{Y}\sigma = 0$ for all Y, and σ is constant by Lemma 5.5. The normalization (5.12) is possible, since $c \mapsto \int e^{-cd_k}$ is constant on $(0,\infty)$ and tends to a nonzero multiple of the body of ρ as $c \to \infty$; we prove the two statements in Section 5.4. We emphasize that $\mathcal{B}_k$ is bilinear, not sesquilinear, and has no positivity property. The weighted L^2 norms below are Hilbert norms, unrelated to $\mathcal{B}_k$.

Since $\widehat{Y}$ is a derivation and $\int \widehat{Y}(ff') = 0$, we have the integration by parts formula

$$\int (\widehat{Y} f) f' = -(-1)^{|Y||f|} \int f\, \widehat{Y} f' \tag{5.13}$$

for homogeneous Y and f and smooth f', provided at least one of f and f' has compact support.

We now define the Casimir operator. For i, j in the index set of Section 5.2, we let $E_{ij} \in \mathfrak{g}$ denote the supermatrix with entry one at (i,j) and zero elsewhere, which is homogeneous of parity $|i| + |j|$.

Definition 5.7. We define the *Casimir operator* to be

$$\mathsf{C}_k = -\frac{1}{2} \sum_{i,j} (-1)^{|j|}\, \widehat{E_{ji}}\, \widehat{E_{ij}}. \tag{5.14}$$

The definition comes from the quadratic Casimir element of $\mathfrak{gl}(2k|2k)$, and we now explain the connection. By Lemma 5.4(d), the operators $\phi(Y) = -\widehat{Y}$ for even Y and $\phi(Y) = \mathrm{i}\widehat{Y}$ for odd Y satisfy

$$[\phi(Y), \phi(Y')] = \phi([Y, Y']) \tag{5.15}$$

for homogeneous Y and Y', which one checks separately in the three cases of two even, one even and one odd, and two odd elements. In other words, ϕ is a representation of the Lie superalgebra $\mathfrak{g}$ by operators on superfunctions. For the bilinear form $\mathrm{STr}(YY')$, the basis $((-1)^{|i|}E_{ji})$ is dual to (E_{ij}), and the quadratic Casimir element of $\mathfrak{gl}(2k|2k)$ is defined as the sum $\sum_{i,j}(-1)^{|i|}E_{ji}E_{ij}$ [8]. Its image under ϕ is a multiple of C_k. Indeed, $\phi(E_{ji})\phi(E_{ij})$ equals $\widehat{E_{ji}}\widehat{E_{ij}}$ if E_{ij} is even and $-\widehat{E_{ji}}\widehat{E_{ij}}$ if E_{ij} is odd, while $(-1)^{|i|} = (-1)^{|j|}$ in the first case and $(-1)^{|i|} = -(-1)^{|j|}$ in the second, and therefore

$$\sum_{i,j} (-1)^{|i|} \phi(E_{ji})\phi(E_{ij}) = \sum_{i,j} (-1)^{|j|} \widehat{E_{ji}}\widehat{E_{ij}} = -2\mathsf{C}_k. \tag{5.16}$$

We will need the following properties of C_k.

Lemma 5.8. *(a) For all homogeneous $Y \in \mathfrak{g}$, we have $\mathsf{C}_k \widehat{Y} = \widehat{Y}\mathsf{C}_k$, and $L_{g'}\mathsf{C}_k = \mathsf{C}_k L_{g'}$ for all $g' \in G_0$. Moreover, $\mathsf{C}_k 1 = 0$.*

(b) For all smooth superfunctions f and f', at least one of which has compact support, we have $\int (\mathsf{C}_k f) f' = \int f\, \mathsf{C}_k f'$.

(c) *Let* s *be a section over a coordinate patch* $\mathcal{O} \subset \mathcal{M}_k^0$. *In the coordinates* (x, χ) *given by* s, *the operator* C_k *is a second-order differential operator of the form*

$$\mathsf{C}_k = -\Delta_{\mathcal{M}_k^0} + \sum_{\mu,\nu} n^{\mu\nu} \partial_{x_\mu} \partial_{x_\nu} + \sum_{\mu} b^\mu \partial_{x_\mu} + c, \tag{5.17}$$

where each $n^{\mu\nu}$ *is the multiplication by a smooth superfunction whose coefficients have degree at least two in the* χ_a, *and* b^μ *and* c *are linear maps of* $\bigwedge_{4k^2}$ *(they may contain* χ_a *and* ∂_{χ_a}*) with smooth coefficients in* x. *Moreover,* C_k *is a finite sum of products of two vector fields in the coordinates* (x, χ). *In particular, it is a finite sum of products of at most two of the derivatives* ∂_{x_μ} *and* ∂_{χ_a}, *with smooth superfunctions as coefficients on the left.*

Proof. We begin with (a). Its first assertion is the statement that the Casimir element is central [8]; for Lie superalgebras in general, see also [132]. We give the short direct proof. We write $c_{ij} = \phi(E_{ij})$, we recall that E_{ij} has the parity $|E_{ij}| = |i| + |j|$ modulo two, and we set

$$\Omega = \sum_{i,j} (-1)^{|i|} c_{ji} c_{ij},$$

so that $\Omega = -2\mathsf{C}_k$ by (5.16). The brackets of the E_{ij}, and (5.15), give

$$[E_{mn}, E_{ji}] = \delta_{nj} E_{mi} - (-1)^{|E_{mn}||E_{ji}|} \delta_{im} E_{jn},$$
$$[c_{mn}, c_{ji}] = \delta_{nj} c_{mi} - (-1)^{|E_{mn}||E_{ji}|} \delta_{im} c_{jn}.$$

Using $[c, c'c''] = [c, c']c'' + (-1)^{|c||c'|} c'[c, c'']$ and $|E_{ji}| = |E_{ij}|$, we find

$$\begin{aligned}
[c_{mn}, \Omega] &= \sum_{i,j} (-1)^{|i|} \Big([c_{mn}, c_{ji}]\, c_{ij} + (-1)^{|E_{mn}||E_{ij}|} c_{ji}\, [c_{mn}, c_{ij}] \Big) \\
&= \sum_i (-1)^{|i|} c_{mi} c_{in} - \sum_j (-1)^{|m| + |E_{mn}||E_{jm}|} c_{jn} c_{mj} \\
&\quad + \sum_j (-1)^{|n| + |E_{mn}||E_{jn}|} c_{jn} c_{mj} - \sum_i (-1)^{|i|} c_{mi} c_{in}.
\end{aligned}$$

The first and the last sums cancel. In the two middle sums, the difference of the exponents is $(|n| - |m|)(1 + |E_{mn}|)$, which is even since $|E_{mn}| = 1$ whenever $|n| \neq |m|$, and these sums cancel as well. Therefore Ω supercommutes with every c_{mn}. Since Ω is even and c_{mn} is a nonzero multiple of $\widehat{E_{mn}}$, we obtain $\mathsf{C}_k \widehat{Y} = \widehat{Y} \mathsf{C}_k$ for all homogeneous Y. The second assertion follows from $L_{g'} \widehat{Y} L_{g'}^{-1} = \widehat{g' Y g'^{-1}}$ and the invariance of the form $\mathrm{STr}(YY')$ under conjugation by g', which maps dual bases to dual bases. The identity $\mathsf{C}_k 1 = 0$ holds since $\widehat{Y} 1 = 0$.

We now prove (b). By linearity, we may assume that f is homogeneous. Let $a = |i| + |j|$. Differentiation does not enlarge support, so the compactly supported factor remains compactly supported in each of the two applications of (5.13). We find

$$\int (\widehat{E_{ji}} \widehat{E_{ij}} f) f' = -(-1)^{a(|f|+a)} \int (\widehat{E_{ij}} f)\, \widehat{E_{ji}} f' = (-1)^a \int f\, \widehat{E_{ij}} \widehat{E_{ji}} f'.$$

We multiply by $(-1)^{|j|}$, sum over i and j, and exchange the names of i and j. Since $(-1)^{|j|+a} = (-1)^{|i|}$, we obtain (b).

It remains to prove (c). We split the sum (5.14) into the terms with E_{ij} even and those with E_{ij} odd. Let Y be in the Lie algebra of G_0. In the coordinates of s, we have $e^{sY}\mathsf{s}(x) = \mathsf{s}(x_s)\mathsf{k}_s$ with $x_s = e^{sY}xe^{-sY}$ and $\mathsf{k}_s \in K_0$. By Lemma 5.4(b) and the equivariance of superfunctions, $\widehat{Y}$ is therefore the sum of the vector field V_Y of the action of e^{sY} on $\mathcal{M}_k^0$ and of the infinitesimal rotation of the odd coordinates by k_s, which has the form $\sum c_{ab}(x)\chi_a\partial_{\chi_b}$ and contains no derivative in x. By complex linearity, the same holds for all $Y \in \mathfrak{g}_0$, with V_Y complex linear in Y.

We now compute the part of the even terms of second order in x. It is the same in all charts obtained from one chart at Λ by the action of an element of G_0, which we call translated charts in Section 5.2, by (a). It therefore suffices to compute it at the point Λ. Let $\mathfrak{m}_\mathbb{R}$ denote the real subspace of $\mathfrak{m}_0$ described in Section 5.2, on which the form $\mathrm{STr}(YY')$ is real and positive definite, and let $Y_1, \dots, Y_{4k^2}$ be an orthonormal basis of $\mathfrak{m}_\mathbb{R}$ for this form. Let $Z_1, \dots, Z_m$ be a basis of $\mathfrak{k}_0$, and let $Z^1, \dots, Z^m$ be its dual basis for $\mathrm{STr}(YY')$. The decomposition $\mathfrak{g}_0 = \mathfrak{m}_0 \oplus \mathfrak{k}_0$ is orthogonal for this form, by the cyclicity of the supertrace and $\Lambda Y = -Y\Lambda$ for $Y \in \mathfrak{m}_0$. The element $\sum(-1)^{|i|}E_{ji} \otimes E_{ij}$ of $\mathfrak{g}_0 \otimes \mathfrak{g}_0$, summed over the even E_{ij}, is the element dual to the form $\mathrm{STr}(YY')$ on $\mathfrak{g}_0$. It can be computed with any basis of $\mathfrak{g}_0$ and its dual basis, and therefore

$$\sum_{i,j:\ |i|=|j|} (-1)^{|j|}\,\widehat{E_{ji}}\,\widehat{E_{ij}} = \sum_\mu \widehat{Y_\mu}\widehat{Y_\mu} + \sum_\nu \widehat{Z^\nu}\widehat{Z_\nu}.$$

The part of second order in x of a product of two of the operators $\widehat{Y}$ at Λ is the symmetrized product of the two tangent vectors $V_Y(\Lambda)$. For $Z \in \mathfrak{k}_0$, the vector $V_Z(\Lambda)$ vanishes, since e^{sZ} commutes with Λ. For $Y \in \mathfrak{m}_\mathbb{R}$, we have $V_Y(\Lambda) = [Y, \Lambda]$, and

$$-\tfrac{1}{2}\,\mathrm{STr}\,\big([Y,\Lambda][Y',\Lambda]\big) = 2\,\mathrm{STr}(YY'), \qquad Y, Y' \in \mathfrak{m}_\mathbb{R},$$

so that the vectors $2^{-1/2}V_{Y_\mu}(\Lambda)$ form an orthonormal basis of $T_\Lambda\mathcal{M}_k^0$ for the metric (5.2). The second-order part at Λ of the right side of the last display is therefore twice that of $\Delta_{\mathcal{M}_k^0}$, and the even terms of (5.14), which are $-\frac{1}{2}$ times this display, have the second-order part of $-\Delta_{\mathcal{M}_k^0}$. This is the classical fact that the Casimir operator acts on a Riemannian symmetric space as a multiple of the Laplace–Beltrami operator (compare [66]). The remaining parts of the even terms have order at most one in x.

For odd Y, the construction of Lemma 5.4 shows that $\widehat{Y}$ is the sum of the left-invariant vector field in the direction A and of a vector field in the odd coordinates. Here A has coefficients of odd degree, since $\vartheta_Y A$ is even. A product of two such fields contains second derivatives in x only with coefficients of degree at least two, which completes the proof of (5.17).

It remains to prove the last assertion of (c). We showed above that, in the coordinates of s, every $\widehat{Y}$ is a vector field in (x, χ) with smooth superfunctions as coefficients. For

even Y, the field is the sum of V_Y and a rotation of the odd coordinates. For odd Y, it is the sum of the left-invariant field in the direction A and a vector field in the odd coordinates, and the left-invariant field is the sum of a vector field in x and a rotation of the odd coordinates, by the argument for even Y with right translation by e^{sA} in place of left translation by e^{sY}. Hence C_k is a finite sum of products of two such vector fields. For vector fields V and V' in the coordinates $z = (x, \chi)$, the Leibniz rule gives

$$VV' = \sum_{I,J} \pm v^I v'^J \partial_{z_I} \partial_{z_J} + \sum_{I,J} v^I (\partial_{z_I} v'^J) \partial_{z_J}, \qquad V = \sum_I v^I \partial_{z_I}, \quad V' = \sum_J v'^J \partial_{z_J},$$

which is of the stated form. In particular, the coefficient of $\partial_{x_\mu} \partial_{x_\nu}$ in C_k is the multiplication by a superfunction. Together with the previous paragraph, this proves the assertion on $n^{\mu\nu}$. □

Remark 5.9. At rank one, we show in Lemma 6.8 that on superfunctions invariant under $\mathfrak{k}$, which are functions of (B, F) in polar coordinates, C_1 acts as $-d\Delta_D d^{-1}$. The term $-\frac{2(B^2-1)}{d}\partial_B + \frac{2(1-F^2)}{d}\partial_F$ of $d\Delta_D d^{-1}$, which is absent from the Laplace–Beltrami operator of the underlying manifold, comes from the odd directions. (This term is the reason why the singlet diffusion of Remark 1.5 differs from the diffusion with generator $X\Delta_D$ of Definition 1.4.)

Remark 5.10. The normalization of C_k is fixed by its body part $-\Delta_{\mathcal{M}_k^0}$ (Lemma 5.8(c)), and in this sense C_k is minus the invariant Laplace operator of $\mathcal{M}_k$. This also holds with the odd directions included. We show in Lemma 16.5, where Part IV uses it, that C_k is minus the Laplace–Beltrami operator of the invariant supermetric $-\frac{1}{2}\operatorname{STr}(\mathrm{d}Q\,\mathrm{d}Q)$ (compare [139]), in the weak sense in which the quadratic form of C_k for the invariant Berezin integral is given by the inverse of this supermetric. The chain of Section 4 produces C_k with this normalization: one link of the chain acts on the tangent variables as the heat semigroup of $-\mathsf{C}_k$ at time $(4s_E^2 W^2)^{-1}$, as explained in Section 5.1, and Lemma 18.8 proves this.

We will need two computations with C_k, which determine its radial part at rank one in Lemma 6.8. For even superfunctions u and v, we set

$$\mathcal{Q}_{\mathsf{C}_k}(u, v) = \mathsf{C}_k(uv) - (\mathsf{C}_k u)v - u\,\mathsf{C}_k v.$$

Lemma 5.11. *(a) For even superfunctions u and v, we have*

$$\mathcal{Q}_{\mathsf{C}_k}(u, v) = -\sum_{i,j} (-1)^{|j|} \big(\widehat{E_{ji}} u\big)\big(\widehat{E_{ij}} v\big), \tag{5.18}$$

and $\mathcal{Q}_{\mathsf{C}_k}(u, v)_\varnothing = -2\nabla u_\varnothing \cdot \nabla v_\varnothing$ for the metric (5.2).

(b) We have $\mathsf{C}_k Q = 0$ *entrywise, and*

$$\mathsf{C}_k d_k = 0, \qquad \mathsf{C}_k \operatorname{STr}\big((\Lambda Q)^2\big) = -8 d_k^2. \tag{5.19}$$

Proof. We begin with (a). Applying the Leibniz rule twice to (5.14), and exchanging i and j in one of the two mixed sums, which is possible since $(-1)^{|i|+|E_{ij}|} = (-1)^{|j|}$, we obtain (5.18). For odd E_{ij}, both factors in (5.18) are odd and their product has no body. For even E_{ij}, the field $\widehat{E_{ij}}$ is the sum of the vector field $V_{E_{ij}}$ on $\mathcal{M}_k^0$ and a rotation of the odd coordinates, as shown in the proof of Lemma 5.8(c), so the body of $\widehat{E_{ij}}u$ is $V_{E_{ij}}u_\varnothing$. The computation of the second-order part in that proof then shows that $\mathcal{Q}_{\mathsf{C}_k}(u,v)_\varnothing = -2\nabla u_\varnothing \cdot \nabla v_\varnothing$.

We now prove (b). Fix i and j, let ϑ and ϑ' be auxiliary parameters of the parity $a = |E_{ij}|$, and set $D = \vartheta\widehat{E_{ij}}$ and $D' = \vartheta'\widehat{E_{ji}}$. These are even derivations, which commute with multiplication by ϑ and ϑ', and by (5.6), we have $DQ = [P,Q]$ and $D'Q = [P',Q]$ with the even supermatrices $P = \vartheta E_{ij}$ and $P' = \vartheta' E_{ji}$. Since $D'D = (-1)^a \vartheta'\vartheta\, \widehat{E_{ji}}\widehat{E_{ij}}$ and $(-1)^{|j|+a} = (-1)^{|i|}$, the definition (5.14) becomes

$$-2\mathsf{C}_k f = \sum_{i,j} (-1)^{|i|}\, [D'Df]_{\vartheta'\vartheta},$$

where $[\,\cdot\,]_{\vartheta'\vartheta}$ denotes the coefficient of $\vartheta'\vartheta$.

We have $D'DQ = [P,[P',Q]]$, and multiplying out with $PP' = \vartheta\vartheta' E_{ii}$, $PQP' = \vartheta\vartheta' Q_{jj}E_{ii}$, $P'QP = \vartheta'\vartheta Q_{ii}E_{jj}$, and $P'P = \vartheta'\vartheta E_{jj}$, we find

$$[D'DQ]_{\vartheta'\vartheta} = (-1)^a\big(E_{ii}Q - Q_{jj}E_{ii}\big) - Q_{ii}E_{jj} + QE_{jj}.$$

We sum over i and j with the sign $(-1)^{|i|}$. Each of the four terms is then a multiple of $\sum_i (-1)^{|i|} = 0$ or of $\operatorname{STr} Q = 0$, by (5.5), so that $\mathsf{C}_k Q = 0$ entrywise. Since d_k is a linear combination of entries of Q, this implies the first identity in (5.19).

For the second identity, the Leibniz rule gives

$$D'D \operatorname{STr}(\Lambda Q \Lambda Q) = 2\operatorname{STr}\big(\Lambda(D'DQ)\Lambda Q\big) + 2\operatorname{STr}\big(\Lambda[P,Q]\Lambda[P',Q]\big),$$

where we used the cyclicity of the supertrace for even supermatrices. After the signed sum, the first term gives $-4\operatorname{STr}(\Lambda(\mathsf{C}_k Q)\Lambda Q) = 0$. For even supermatrices M and N, a direct computation gives $\operatorname{STr}(MPNP') = (-1)^{|j|}\vartheta'\vartheta\, M_{ii}N_{jj}$. Expanding the commutators and using cyclicity, we obtain, with $\Lambda = \operatorname{diag}(\lambda_1,\dots,\lambda_{4k})$,

$$\big[\operatorname{STr}\big(\Lambda[P,Q]\Lambda[P',Q]\big)\big]_{\vartheta'\vartheta} = (-1)^{|j|}\Big(2\lambda_i\lambda_j Q_{ii}Q_{jj} - \lambda_i(Q\Lambda Q)_{jj} - \lambda_j(Q\Lambda Q)_{ii}\Big).$$

The signed sum of the first term is $2(\operatorname{STr}\Lambda Q)^2 = 8d_k^2$, and those of the other two terms are multiples of $\operatorname{STr}\Lambda = 0$. Therefore $-2\mathsf{C}_k \operatorname{STr}((\Lambda Q)^2) = 2\cdot 8d_k^2$, which is the second identity in (5.19). □

Coefficient norms and weighted spaces

We now define the norms in which we measure the coefficients of superfunctions, here and in Part IV, and the weighted spaces on which the semigroup of Proposition 5.15 acts. Part III does not use these norms, except through the spaces L^2_p and the multipliers defined below, and its reader may skip Definition 5.12 and Lemma 5.13. The norms are defined for the Grassmann algebra $\bigwedge_r$ of Definition 3.1 with any number r of generators $\xi_1, \dots, \xi_r$, and they measure the odd coordinates on a scale that may differ from one. Let $\omega = (\omega_1, \dots, \omega_r) \in (0,\infty)^r$ be a weight vector, and set $\omega^I = \prod_{a\in I} \omega_a$.

Definition 5.12. For $f \in \bigwedge_r$, we define

$$|f|_{1;\omega} = \sum_{I\subset[r]} \omega^I |f_I|, \qquad |f|_{2;\omega} = \Big(\sum_{I\subset[r]} \omega^{2I} |f_I|^2 \Big)^{1/2}.$$

We write $|f|_1$ and $|f|_2$ when $\omega_a = 1$ for all a, and $|f|_{1;\omega}$ and $|f|_{2;\omega}$ with a number $\omega > 0$ when $\omega_a = \omega$ for all a. For a matrix $M = \sum_I M_I \xi^I$ with entries in $\bigwedge_r$, where the M_I are complex matrices, we define $\|M\|_{1;\omega} = \sum_I \omega^I \|M_I\|_{\mathrm{op}}$, where $\|\cdot\|_{\mathrm{op}}$ is the operator norm. For superfunctions on a superdomain (Definition 3.2), these norms are taken pointwise.

Equivalently, $|f|_{1;\omega}$ and $|f|_{2;\omega}$ are the ℓ^1 and ℓ^2 norms of the coefficients of f in the normalized monomials $\omega^{-I}\xi^I$. If we write $\xi_a = \omega_a \hat\xi_a$, then $|f|_{1;\omega}$ is the norm $|\cdot|_1$ in the monomials $\hat\xi^I$. In Part IV, the weight is the number $\omega = h = W^{-1/2}$, and we call $h^{-|I|}\xi^I$ the *h-scaled monomials*. A coefficient of degree s then carries the factor h^s, and a derivative in an odd coordinate costs a factor h^{-1}, as the following lemma makes precise.

Lemma 5.13. *Let $\omega \in (0,\infty)^r$, let $f, f' \in \bigwedge_r$, and let M and M' be matrices with entries in $\bigwedge_r$ such that MM' is defined.*

(a) We have $|ff'|_{1;\omega} \le |f|_{1;\omega}|f'|_{1;\omega}$, $|ff'|_{2;\omega} \le |f|_{1;\omega}|f'|_{2;\omega}$, and $\|MM'\|_{1;\omega} \le \|M\|_{1;\omega}\|M'\|_{1;\omega}$.

(b) For all $I \subset [r]$, multiplication by $\omega^{-I}\xi^I$ is a partial isometry for $|\cdot|_{2;\omega}$, that is, it is isometric on the span of the monomials ξ^J with $J \cap I = \emptyset$, and it annihilates the other monomials.

(c) For all $a \in [r]$ and $i \in \{1,2\}$, we have $|\partial_{\xi_a} f|_{i;\omega} \le \omega_a^{-1} |f|_{i;\omega}$.

(d) Suppose that $\omega_a \le \omega_* \le 1$ for all a, and that $f_I = 0$ for all I with $|I| < s$. Then $|f|_{i;\omega} \le \omega_*^s |f|_i$ for $i \in \{1,2\}$.

(e) Let $a \in \bigwedge_r$ be even with soul n, and let g be holomorphic in a neighborhood of $a_\emptyset$. Then

$$|g(a)|_{1;\omega} \le \sum_{j=0}^{\lfloor r/2 \rfloor} \frac{|g^{(j)}(a_\emptyset)|}{j!} |n|_{1;\omega}^j,$$

and in particular $|e^a|_{1;\omega} \le e^{\mathrm{Re}\, a_\emptyset + |n|_{1;\omega}}$.

(f) Let $B \in \mathbb{C}^{r\times r}$ be invertible, let $\xi'_a = \sum_b B_{ab}\xi_b$, and let $|f|'_1$ denote the norm $|\cdot|_1$ in the monomials of $\xi'_1, \dots, \xi'_r$. Then $|f|_1 \le \max(1, \|B\|_\infty)^r |f|'_1$, where $\|B\|_\infty = \max_a \sum_b |B_{ab}|$.

Proof. We begin with (b). For disjoint I and J, we have $\xi^I\xi^J = \pm\xi^{I\cup J}$ and $\omega^{I\cup J} = \omega^I\omega^J$, while $\xi^I\xi^J = 0$ otherwise. These relations imply (b) and the bound $|\xi^I f'|_{i;\omega} \le \omega^I |f'|_{i;\omega}$ for $i \in \{1,2\}$. Expanding $f = \sum_I f_I\xi^I$ and using the triangle inequality, we find

$$|ff'|_{i;\omega} \le \sum_I |f_I|\,|\xi^I f'|_{i;\omega} \le |f|_{1;\omega}|f'|_{i;\omega}, \qquad i \in \{1,2\},$$

and the bound for matrices follows in the same way from $\|M_I M'_J\|_{\mathrm{op}} \le \|M_I\|_{\mathrm{op}}\|M'_J\|_{\mathrm{op}}$. This proves (a).

Next, for $a \in J$ we have

$$\partial_{\xi_a}\xi^J = \pm\xi^{J\setminus\{a\}}, \qquad \omega^{J\setminus\{a\}} = \omega_a^{-1}\omega^J,$$

while $\partial_{\xi_a}\xi^J = 0$ for $a \notin J$, and (c) follows. Statement (d) follows from $\omega^I \le \omega_*^{|I|} \le \omega_*^s$ for $|I| \ge s$. By (3.6) and (a), we have

$$|g(a)|_{1;\omega} \le \sum_{j=0}^{\lfloor r/2\rfloor} \frac{|g^{(j)}(a_\varnothing)|}{j!}\,|n^j|_{1;\omega} \le \sum_{j=0}^{\lfloor r/2\rfloor} \frac{|g^{(j)}(a_\varnothing)|}{j!}\,|n|_{1;\omega}^j,$$

which is (e), and $|g^{(j)}(a_\varnothing)| = e^{\operatorname{Re} a_\varnothing}$ for $g = \exp$.

It remains to prove (f). We write $f = \sum_I f'_I\xi'^I$. By (a), we have

$$|\xi'^I|_1 \le \prod_{a\in I} |\xi'_a|_1 \le \|B\|_\infty^{|I|},$$

and the claim follows from the triangle inequality. □

We measure superfunctions on $\mathcal{M}_k$ in these norms with a scalar weight $\omega \in (0,1]$, which will be chosen in the proof of Proposition 5.15. Recall from Section 5.2 that two sections differ by a unitary change of odd coordinates. The norm $|f|_{2;\omega}$ is therefore the same in all sections and defines a function on $\mathcal{M}_k^0$. In the norm $|\cdot|_{2;\omega}$, a coefficient of odd degree j carries the factor ω^j. We call $|\cdot|_{2;\omega}$ the *degree metric* with parameter ω; Part IV also uses it, with the parameters ω and $\varkappa_k$. For a superfunction f and $p \in \mathbb{R}$, we set

$$\|f\|_p^2 = \int_{\mathcal{M}_k^0} v^{2p}|f|_{2;\omega}^2 \,\mathrm{dvol}, \tag{5.20}$$

and we let L_p^2 denote the Hilbert space completion of the compactly supported superfunctions in this norm. Its elements are measurable maps $f\colon G_0 \to \bigwedge_{4k^2}$ with the equivariance of Definition 5.2 and $\|f\|_p < \infty$. The norms for different $\omega \in (0,1]$ are equivalent, with

constants depending on ω and k. For $q \geq p$, we have $L^2_q \subset L^2_p$, since $v \geq 1$. Writing $f = \sum_I f_I \chi^I$ and $f' = \sum_I f'_I \chi^I$ in the coordinates of a section, we let

$$\langle f, f' \rangle_p = \int_{\mathcal{M}^0_k} v^{2p} \sum_I \omega^{2|I|} f_I \, \overline{f'_I} \, \mathrm{dvol}$$

denote the inner product of L^2_p, in which the complex conjugate is taken coefficientwise. A change of section acts on the coefficients of each degree by a unitary matrix, so that this inner product is the same in all sections. It is unrelated to $\mathcal{B}_k$.

Derivatives of superfunctions are taken in charts adapted to the action of G_0. Let $\mathfrak{m}_{\mathbb{R}} = \mathfrak{m}_0 \cap \mathrm{Lie}(G_0)$, with the inner product $2\,\mathrm{STr}(XX')$ induced by the metric (5.2) under $X \mapsto [X, \Lambda]$. We choose a sufficiently small ball about zero in $\mathfrak{m}_{\mathbb{R}}$, let $\mathcal{O}_0$ be its image under $X \mapsto e^X \Lambda e^{-X}$, and use the coordinates of X in an orthonormal basis as normal coordinates. The ball and $\mathcal{O}_0$ are K_0-invariant, and we fix the exponential section

$$\mathsf{s}_0(e^X \Lambda e^{-X}) = e^X.$$

For $g' \in G_0$, we call the map $x \mapsto g' x g'^{-1}$ on $\mathcal{O}_0$, together with the section $\mathsf{s}_{g'}(y) = g' \, \mathsf{s}_0(g'^{-1} y g')$ over $g' \mathcal{O}_0 g'^{-1}$, a *translated chart*. In a translated chart, a superfunction f becomes $f_{g'}(X, \chi) = f(g' e^X; \Theta)$, and we let $|\nabla^j f|_{2;\omega}(g' \Lambda g'^{-1})$ denote the norm $|\cdot|_{2;\omega}$ of the collection of j-th partial derivatives in X of its coefficients, at $X = 0$. The exponential section satisfies $\mathsf{s}_0(\mathsf{k} x \mathsf{k}^{-1}) = \mathsf{k} \mathsf{s}_0(x) \mathsf{k}^{-1}$ for $\mathsf{k} \in K_0$, so equivariance gives

$$f_{g'\mathsf{k}}(X, \chi) = \alpha_{\mathsf{k}}\big(f_{g'}(\mathrm{Ad}_{\mathsf{k}} X, \chi)\big).$$

Thus two translated charts centered at the same point differ by an orthogonal linear change of normal coordinates and a constant unitary change of odd coordinates. Derivatives of order j transform only among derivatives of that order, and their norms differ by at most a factor depending on j and k. We say that a superfunction m is a *multiplier of degree* $s \geq 0$ if, for every integer $j \geq 0$, there exists C_j such that $|\nabla^j m|_{1;1} \leq C_j v^s$ on $\mathcal{M}^0_k$. By Lemma 5.13(a), multiplication by a multiplier of degree s maps L^2_p to L^2_{p-s} boundedly, for all p.

The operators of the sigma model look the same in all translated charts. By Lemma 5.8(a), the expression (5.17) of C_k in the chart of g' is the same for all g', and the invariant integral (5.11) has the same density ρ in all of them. In this sense, $\mathcal{M}_k$ has bounded geometry: all coefficients of C_k and of the integral, and all their derivatives, are bounded uniformly over the translated charts.

Lemma 5.14. *(a) The entries of Q^{BB} are multipliers of degree one, the entries of Q^{BF} and Q^{FB} are multipliers of degree $\frac{1}{2}$, and the entries of Q^{FF} are multipliers of degree zero.*

(b) The soul of d_k has coefficients of degree at least two in the χ_a, and $|d_k - (d_k)_\emptyset|_{1;\omega} \leq C\omega^2 v$ for all $\omega \in (0, 1]$.

Proof. By the polar decomposition of $U(k,k)$, we may represent the point $g\Lambda g^{-1}$ with $g = \operatorname{diag}(g_{\mathrm{B}}|g_{\mathrm{F}})$ and g_{B} positive definite. The proof of Lemma 5.1(a) then yields

$$\|g_{\mathrm{B}}\|_{\mathrm{op}}^2 = \|g_{\mathrm{B}}^{-1}\|_{\mathrm{op}}^2 \le \operatorname{Tr}(g_{\mathrm{B}} g_{\mathrm{B}}^*) = 2(v-1),$$

and g_{F} is unitary. The entries of $\Lambda e^{-2\Theta}$ are polynomials in the χ_a with bounded coefficients, and $Q = g\Lambda e^{-2\Theta} g^{-1}$. The blocks Q^{BB}, Q^{BF}, and Q^{FB}, and Q^{FF} are therefore bounded by $C\|g_{\mathrm{B}}\|_{\mathrm{op}}^2$, $C\|g_{\mathrm{B}}\|_{\mathrm{op}}$, and C, respectively, and the same bounds for the derivatives follow from the argument of Lemma 5.1(b). For (b), we write

$$d_k = \tfrac{1}{2} \operatorname{STr}(\Lambda g \Lambda e^{-2\Theta} g^{-1}).$$

The terms of odd degree in Θ vanish, since g is block diagonal in the gradings and odd powers of Θ have vanishing diagonal blocks, and by Lemma 5.13(d) the terms of degree at least two are bounded by $C\omega^2 v$ in $|\cdot|_{1;\omega}$. □

The entries of Q^{FF}, including their souls, are bounded because $Q^{\mathrm{FF}} = g_{\mathrm{F}} (\Lambda e^{-2\Theta})^{\mathrm{FF}} g_{\mathrm{F}}^{-1}$ by (5.4), which involves the compact factor g_{F} only. This is why the fermionic sources in the next section are bounded perturbations of the generator.

Finally, we fix cutoffs adapted to the weight v. Let $\psi \in C_c^\infty(\mathbb{R})$ be equal to one on $[-1,1]$ and to zero outside $[-2,2]$, and set

$$\psi_R = \psi(R^{-1} \log v), \qquad R \ge 1. \tag{5.21}$$

Let r_{B} denote the distance from Λ_{B} to Q^{BB} in H_k. By the proof of Lemma 5.1(c), we have $v \ge \cosh(r_{\mathrm{B}}/\sqrt{k})$. Thus ψ_R has compact support: its support has bounded r_{B}, and S_k is compact. Moreover, $|\nabla \psi_R| + |\nabla^2 \psi_R| \le C R^{-1}$ by Lemma 5.1(b).

5.3 The partition function

We now define the partition function. The heuristic origin of its generator in the chain of Section 4 is explained in Section 5.1. Throughout this section, $X > 0$ and $\varepsilon > 0$. Recall from Section 4 the source supermatrix $\Gamma = \operatorname{diag}(\varepsilon\Lambda_{\mathrm{B}}|\Gamma_{\mathrm{F}})$, with

$$\Gamma_{\mathrm{F}} = \operatorname{diag}(b_1 - \mathrm{i}u_1, \dots, b_k - \mathrm{i}u_k, -b_1 - \mathrm{i}u_1, \dots, -b_k - \mathrm{i}u_k).$$

In this section we allow Γ_{F} to be an arbitrary complex diagonal matrix. Using

$$2\pi\varepsilon d_k(Q) = \pi\varepsilon \operatorname{Tr}(\Lambda_{\mathrm{B}} Q^{\mathrm{BB}}) - \pi\varepsilon \operatorname{Tr}(\Lambda_{\mathrm{F}} Q^{\mathrm{FF}}),$$

we find

$$\pi \operatorname{STr}(Q\Gamma) = 2\pi\varepsilon\, d_k(Q) - \pi \operatorname{Tr}\big(Q^{\mathrm{FF}}(\Gamma_{\mathrm{F}} - \varepsilon\Lambda_{\mathrm{F}})\big). \tag{5.22}$$

The first term is the damping. The second term carries the sources, and it vanishes at the base point $(\mathbf{u}, \mathbf{b}) = (0, \varepsilon\mathbf{1})$, where $\Gamma = \Gamma_0 = \varepsilon\Lambda$. We define the *generator* to be the differential operator

$$\mathsf{G}_{X,\Gamma} = X\mathsf{C}_k + \pi\,\mathrm{STr}(Q\Gamma), \tag{5.23}$$

acting on superfunctions, where the second term acts by multiplication. For $z \geq 0$, we also set $\mathsf{G}_{X,z} = X\mathsf{C}_k + zd_k$, so that $\mathsf{G}_{X,\Gamma_0} = \mathsf{G}_{X,2\pi\varepsilon}$. When the parameters are fixed, we write $\mathsf{G} = \mathsf{G}_{X,\Gamma}$. With the semigroup written as $e^{-t\mathsf{G}}$, the body part $-\Delta_{\mathcal{M}_k^0}$ of C_k is a nonnegative operator on the underlying manifold, and the damping term $2\pi\varepsilon d_k$ has nonnegative body.

The following proposition constructs the semigroup $e^{-t\mathsf{G}}$ on the weighted spaces L^2_p of Section 5.2. The constant function 1 belongs to L^2_{-q} for large q but not to L^2_0, and this is why we need all the weights at once. Recall that q_1 is the exponent of Lemma 5.1(c). We call a superfunction f *rapidly decreasing* if it is smooth and $\|\nabla^j f\|_p < \infty$ for all $j \geq 0$ and $p \in \mathbb{R}$, where $\|\nabla^j f\|_p$ is defined by (5.20) with $|\nabla^j f|_{2;\omega}$ in place of $|f|_{2;\omega}$.

Proposition 5.15. *Let $X > 0$, $\varepsilon > 0$, $R_{\mathrm{F}} > 0$, and $\Gamma = \mathrm{diag}(\varepsilon\Lambda_{\mathrm{B}}|\Gamma_{\mathrm{F}})$ with $\|\Gamma_{\mathrm{F}}\|_{\mathrm{op}} \leq R_{\mathrm{F}}$. There exists $\omega_0 > 0$, depending on k, X, ε, and R_{F}, such that the following statements hold for the norms of* (5.20) *with $\omega \in (0, \omega_0]$.*

(a) Realization. *For every $p \in \mathbb{R}$, the operator G, defined on the compactly supported smooth superfunctions, is closable in L^2_p. Its closure $\mathsf{G}_{(p)}$ is the operator associated by Kato's first representation theorem with a closed sectorial form, shifted by a constant, whose domain is*

$$V_p = \{f \in L^2_p : \|\nabla f\|_p + \|v^{1/2} f\|_p < \infty\}.$$

In particular $-\mathsf{G}_{(p)}$ generates a holomorphic semigroup $e^{-t\mathsf{G}_{(p)}}$ on L^2_p, and the compactly supported smooth superfunctions are a core for $\mathsf{G}_{(p)}$.

(b) Consistency and smoothing. *For $q \geq p$ and $f \in L^2_q$, we have $e^{-t\mathsf{G}_{(q)}} f = e^{-t\mathsf{G}_{(p)}} f$ for all $t \geq 0$; we therefore write $e^{-t\mathsf{G}}$. For all $p \in \mathbb{R}$, all integers $m \geq 0$, and all $T > 0$, there exists C such that*

$$\|e^{-t\mathsf{G}}\|_{L^2_p \to L^2_{p+m/2}} \leq Ct^{-m/2}, \qquad 0 < t \leq T.$$

For $t > 0$ and $f \in L^2_p$, the superfunction $e^{-t\mathsf{G}} f$ is rapidly decreasing.

(c) Transpose symmetry. *For all $p \in \mathbb{R}$, the pairing $\mathcal{B}_k$ is bounded on $L^2_p \times L^2_{-p}$, and*

$$\mathcal{B}_k(e^{-t\mathsf{G}} f, f') = \mathcal{B}_k(f, e^{-t\mathsf{G}} f'), \qquad f \in L^2_p,\ f' \in L^2_{-p},\ t \geq 0.$$

(d) The split pairing. *We have $1 \in L^2_{-q}$ for all $q > q_1/2$, and $e^{-t\mathsf{G}}1$ is rapidly decreasing for all $t > 0$. For $0 < a < T$,*

$$\mathcal{B}_k(e^{-a\mathsf{G}}1, e^{-(T-a)\mathsf{G}}1) = \int e^{-T\mathsf{G}}1.$$

In particular, the left side does not depend on a.

(e) Small damping. *Let $z_0 > 0$. For $0 < z \le z_0$, the semigroup $e^{-t\mathsf{G}_{X,z}}$ is the one of (a) with $\Gamma = (z/2\pi)\Lambda$. For $z = 0$, the operator $X\mathsf{C}_k$, defined on the compactly supported smooth superfunctions, is closable in L^2_p, and its closure is the operator associated with a closed sectorial form, shifted by a constant, whose domain is*

$$H^1_p = \{f \in L^2_p : \|\nabla f\|_p < \infty\};$$

we write $e^{-t\mathsf{G}_{X,0}}$ for the semigroup that it generates. For every $p \in \mathbb{R}$, there exists $\omega_0 > 0$ such that, for each fixed $\omega \in (0, \omega_0]$, there is a constant C_ω, depending also on k, X, p, and z_0, with

$$\|e^{-t\mathsf{G}_{X,z}}\|_{L^2_p \to L^2_p} \le e^{C_\omega t}, \qquad 0 \le z \le z_0, \quad t \ge 0.$$

Moreover, let $m_1, \dots, m_n$ be multipliers and $T > 0$. Then the coefficients of the superfunction $e^{-s_n\mathsf{G}_{X,z}} m_n \cdots e^{-s_1\mathsf{G}_{X,z}} m_1 1$, and all their derivatives, are bounded on compact subsets of $\mathcal{M}^0_k$, uniformly in $0 \le z \le z_0$ and $s_1, \dots, s_n \in [0, T]$.

The thresholds ω_0 can be chosen uniformly for (X, ε) in compact subsets of $(0, \infty)^2$. With an admissible $\omega > 0$ fixed, the constants in (b) and (e) can likewise be chosen uniformly on these compact sets; in (e) the bounds remain uniform for $z \in [0, z_0]$. These constants may grow as $\omega \downarrow 0$. The realizations do not depend on $\omega \in (0, \omega_0]$, since the norms for different ω are equivalent.

The statement allows complex fermionic sources, which are needed in Section 8. We prove the proposition in Appendix C, and we now define the partition function.

Definition 5.16. Let $X > 0$, $\varepsilon > 0$, $\mathbf{u} \in \mathbb{R}^k$ with $\sum_j u_j = 0$, and $\mathbf{b} \in (0, \infty)^k$, and let Γ be the source supermatrix of $(\mathbf{u}, \mathbf{b})$. We define the *partition function* of the sigma model by

$$\mathcal{Z}^{(k)}_{X,\varepsilon}(\mathbf{u}, \mathbf{b}) = \mathcal{B}_k\big(e^{-a\mathsf{G}_{X,\Gamma}}1,\, e^{-(1-a)\mathsf{G}_{X,\Gamma}}1\big), \qquad 0 < a < 1, \tag{5.24}$$

where the semigroup is the one of Proposition 5.15.

By Proposition 5.15(d), the right side of (5.24) is independent of a, and

$$\mathcal{Z}^{(k)}_{X,\varepsilon}(\mathbf{u}, \mathbf{b}) = \int e^{-\mathsf{G}_{X,\Gamma}}1.$$

The split form (5.24) is the one that the chain produces in Part IV, where the two ends of the chain are treated separately (see Section 19). By Corollary 5.20 below, we have $\mathcal{Z}^{(k)}_{X,\varepsilon}(0, \varepsilon\mathbf{1}) = 1$.

We now record two further facts about the semigroup, which we use from Section 5.4 on. The proof of Proposition 5.15 also identifies the domains of the realizations. For all $p \in \mathbb{R}$, we have

$$D(\mathsf{G}_{(p)}) = \{f \in V_p : \mathsf{G}f \in L^2_p\}, \tag{5.25}$$

where $\mathsf{G}f$ is taken in the sense of distributions, and $\mathsf{G}_{(p)}f = \mathsf{G}f$ for f in this domain. For an integer $j \geq 1$ and $p \in \mathbb{R}$, we let $D(\mathsf{G}_{(p)}^j)$ denote the set of $f \in L_p^2$ such that $f, \mathsf{G}f, \dots, \mathsf{G}^{j-1}f \in D(\mathsf{G}_{(p)})$, and we set $D(\mathsf{G}_{(p)}^0) = L_p^2$.

Let $X > 0$, $z_0 > 0$, and $T > 0$. We call a superfunction of the form

$$f = e^{-s_n \mathsf{G}_{X,z}} m_n \cdots e^{-s_1 \mathsf{G}_{X,z}} m_1 1, \tag{5.26}$$

where $m_1, \dots, m_n$ are multipliers, $0 \leq z \leq z_0$, and $s_1, \dots, s_n \in [0, T]$, a *semigroup product*. For $z > 0$, Proposition 5.15(b) implies that $e^{-t\mathsf{G}}f$ is rapidly decreasing for $t > 0$, but the bounds degenerate as $t \downarrow 0$. Part III evaluates semigroup products, and their components in the polar coordinates of Section 6, for s_n close to zero (Section 10.7), and the proofs in Sections 6.2 and 6.3 need the regularity of $e^{-t\mathsf{G}}f$ up to $t = 0$. The following lemma provides these bounds, including at $z = 0$. We prove it in Appendix C.6, with the domains of powers of the generator and interior elliptic estimates.

Lemma 5.17. *Let $X > 0$, $z_0 > 0$, and $T > 0$, and fix multipliers $m_1, \dots, m_n$. Let $\mathsf{G} = \mathsf{G}_{X,z}$ with $0 \leq z \leq z_0$, and let f be the semigroup product* (5.26)*. In (a) and (b), the constants do not depend on $z \in [0, z_0]$ or on $s_1, \dots, s_n \in [0, T]$.*

(a) For every integer $j \geq 0$, there exist $p \in \mathbb{R}$ and C such that $f \in D(\mathsf{G}_{(p)}^j)$ and

$$\sum_{l=0}^{j} \|\mathsf{G}^l f\|_p + \sum_{i=0}^{2j} \|\nabla^i f\|_p \leq C. \tag{5.27}$$

(b) For every compact set $K \subset \mathcal{M}_k^0$ and every $j \geq 0$, the coefficients of f and their derivatives of order at most j, in the translated charts, are bounded on K by a constant.

(c) Fix z and the times $s_1, \dots, s_n$. Then the coefficients of $e^{-t\mathsf{G}}f$, in the coordinates of a section, are smooth functions of $t \in [0, \infty)$ and of the body point, with derivatives continuous up to $t = 0$, and $u(t) = e^{-t\mathsf{G}}f$ solves $\partial_t u = -\mathsf{G}u$. For every $j \geq 0$, there exist $p \in \mathbb{R}$ and C such that $\|\nabla^i e^{-t\mathsf{G}}f\|_p \leq C$ for all $t \in [0, T]$ and $i \leq j$.

(d) Statement (c) holds for every rapidly decreasing superfunction f in place of the semigroup product, and then the last bound holds for every $p \in \mathbb{R}$, with C depending on p.

(e) Let $A_{X,z} = -X\Delta_{\mathcal{M}_k^0} + z(d_k)_\emptyset$, acting on complex functions on $\mathcal{M}_k^0$, and use the norms (5.20) *with $|f|$ in place of $|f|_{2;\omega}$. For $z > 0$, the realization and smoothing statements of Proposition 5.15(a) and (b) hold for $A_{X,z}$, with the form domain V_p and the positive-time weight gain stated there. At $z = 0$, use the closure of $-X\Delta_{\mathcal{M}_k^0}$ with form domain H_p^1, as in Proposition 5.15(e). The small-damping statement (e) and statements (a)–(d) above hold for $A_{X,z}$ throughout $0 \leq z \leq z_0$, with multipliers that are smooth functions m on $\mathcal{M}_k^0$ such that $|\nabla^j m| \leq C_j v^s$ for all j.*

In particular, part (b) is the last assertion of Proposition 5.15(e), with last time $s_n \in [0, T]$. The proof of Proposition 5.15 in Appendix C deduces that assertion from the lemma, whose proof uses only the other parts of the proposition.

5.4 Localization and normalization

We now prove the supersymmetric localization identity on $\mathcal{M}_k$. The mechanism is that of Proposition 3.19. An odd vector field $\mathfrak{q}$ with zero divergence, and a $\mathfrak{q}$-exact and $\mathfrak{q}$-closed function whose body has a unique nondegenerate minimum, allow one to insert a Gaussian factor into the integral of a $\mathfrak{q}$-closed function, and then to concentrate the integral at the minimum. Historically, this is the localization principle of Parisi and Sourlas [103], and a general form on compact supermanifolds is due to Schwarz and Zaboronsky [113]. Since $\mathcal{M}_k$ is not compact, their theorem does not apply directly. For the noncompact superspace $H^{2|2}$, the corresponding identity is proved in [28, Appendix C], and for flat superspace in [7, Theorem 11.4.5]; see also Lemma 3.18. We follow these proofs, with the same three steps as in Proposition 3.19, and we control the contributions from infinity by the growth bounds of Section 5.2.

Let $q_{\mathrm{A}} \in \mathfrak{k}_1$ denote the odd supermatrix with the blocks $(q_{\mathrm{A}})^{\mathrm{BF}}_{\mathrm{AA}} = (q_{\mathrm{A}})^{\mathrm{FB}}_{\mathrm{AA}} = \mathrm{Id}_k$ and with all other entries zero, and define q_{R} in the same way with the retarded flavors. We set $\mathfrak{q}_{\mathrm{A}} = \widehat{q_{\mathrm{A}}}$ and $\mathfrak{q}_{\mathrm{R}} = \widehat{q_{\mathrm{R}}}$, and we call them the advanced and retarded *supercharges*. We also define the odd superfunctions

$$\mathfrak{o}_{\mathrm{A}} = -\frac{1}{2}\sum_{i,j}(-1)^{|i|}(q_{\mathrm{A}})_{ij}\,Q_{ji}, \qquad \mathfrak{o}_{\mathrm{R}} = \frac{1}{2}\sum_{i,j}(-1)^{|i|}(q_{\mathrm{R}})_{ij}\,Q_{ji}. \tag{5.28}$$

For an auxiliary odd generator ϑ, we have $\vartheta\mathfrak{o}_{\mathrm{A}} = -\frac{1}{2}\,\mathrm{STr}(\vartheta q_{\mathrm{A}} Q)$, where ϑq_{A} is an even supermatrix.

Lemma 5.18. *We have* $q_{\mathrm{A}}^2 = \mathrm{Id}_{\mathrm{A}}$ *and* $[q_{\mathrm{A}}, \Lambda] = 0$, *and*

$$\mathfrak{q}_{\mathrm{A}} d_k = 0, \qquad \mathfrak{q}_{\mathrm{A}}\mathfrak{o}_{\mathrm{A}} = d_k.$$

The same holds with R *in place of* A. *Moreover, let* $\mathfrak{k}_{\mathrm{A}} = \{Y \in \mathfrak{k} : Y = \mathrm{Id}_{\mathrm{A}} Y \mathrm{Id}_{\mathrm{A}}\}$ *denote the advanced subalgebra, and* $\mathfrak{k}_{\mathrm{R}}$ *the retarded one. Then a superfunction that is invariant under* $\mathfrak{k}_{\mathrm{A}}$ *satisfies* $\mathfrak{q}_{\mathrm{A}} f = 0$, *and the entries of* Q_{RR} *are invariant under* $\mathfrak{k}_{\mathrm{A}}$.

Proof. The first two identities are immediate from the definition of q_{A}. Let ϑ and ϑ' be auxiliary odd generators. Using (5.6) and $[q_{\mathrm{A}}, \Lambda] = 0$, we have

$$\vartheta\mathfrak{q}_{\mathrm{A}} d_k = \tfrac{1}{2}\,\mathrm{STr}([\Lambda, \vartheta q_{\mathrm{A}}]Q) = 0.$$

Using also the cyclicity of the supertrace for the even supermatrices ϑq_{A} and $\vartheta' q_{\mathrm{A}}$, we find

$$\vartheta\mathfrak{q}_{\mathrm{A}}(\vartheta'\mathfrak{o}_{\mathrm{A}}) = -\frac{1}{2}\operatorname{STr}\left(\vartheta' q_{\mathrm{A}}[\vartheta q_{\mathrm{A}}, Q]\right) = -\frac{1}{2}\operatorname{STr}\left([\vartheta' q_{\mathrm{A}}, \vartheta q_{\mathrm{A}}]Q\right) = -\vartheta'\vartheta\operatorname{STr}(q_{\mathrm{A}}^2 Q),$$

where we used $\vartheta' q_{\mathrm{A}}\vartheta q_{\mathrm{A}} = \vartheta'\vartheta q_{\mathrm{A}}^2$ and $\vartheta q_{\mathrm{A}}\vartheta' q_{\mathrm{A}} = -\vartheta'\vartheta q_{\mathrm{A}}^2$. Since $\mathfrak{q}_{\mathrm{A}}$ is odd, we have

$$\vartheta\mathfrak{q}_{\mathrm{A}}(\vartheta'\mathfrak{o}_{\mathrm{A}}) = -\vartheta\vartheta'\mathfrak{q}_{\mathrm{A}}\mathfrak{o}_{\mathrm{A}} = \vartheta'\vartheta\mathfrak{q}_{\mathrm{A}}\mathfrak{o}_{\mathrm{A}},$$

and comparing with the previous display, we obtain

$$\mathfrak{q}_{\mathrm{A}}\mathfrak{o}_{\mathrm{A}} = -\operatorname{STr}(\mathrm{Id}_{\mathrm{A}} Q) = -\operatorname{STr} Q_{\mathrm{AA}}.$$

By (5.5), we have $\operatorname{STr} Q_{\mathrm{RR}} + \operatorname{STr} Q_{\mathrm{AA}} = \operatorname{STr} Q = 0$, and hence

$$d_k = \tfrac{1}{2}\big(\operatorname{STr} Q_{\mathrm{RR}} - \operatorname{STr} Q_{\mathrm{AA}}\big) = -\operatorname{STr} Q_{\mathrm{AA}} = \operatorname{STr} Q_{\mathrm{RR}}.$$

This proves the second identity, and the retarded case is the same. Finally, $q_{\mathrm{A}} \in \mathfrak{k}_{\mathrm{A}}$, and for $Y \in \mathfrak{k}_{\mathrm{A}}$ the commutator $[Y, Q]$ has vanishing flavor block RR. □

We say that a superfunction f has *polynomially integrable coefficients* if $\int v^K |f|_{1;1}\,\mathrm{dvol} < \infty$ for all $K \geq 0$. By the Cauchy–Schwarz inequality and $1 \in L^2_{-q}$, every rapidly decreasing superfunction has this property.

Lemma 5.19. *Let A be a smooth superfunction on $\mathcal{M}_k$ with polynomially integrable coefficients, and suppose that $\mathfrak{q}_{\mathrm{A}} A = 0$ or $\mathfrak{q}_{\mathrm{R}} A = 0$. Then*

$$\int A = A_{\emptyset}(\Lambda).$$

The same holds entrywise for matrices of superfunctions whose entries satisfy these hypotheses.

Proof. We treat $\mathfrak{q} = \mathfrak{q}_{\mathrm{A}}$ and $\mathfrak{o} = \mathfrak{o}_{\mathrm{A}}$; the retarded case is identical. For $c > 0$, we set

$$I(c) = \int A e^{-cd_k}.$$

We will show that I is constant, that $I(c) \to \int A$ as $c \downarrow 0$, and that $I(c) \to A_{\emptyset}(\Lambda)$ as $c \to \infty$. The polynomial integrability of A is used only in the second step; the first and the third use only that A is integrable, because of the factor e^{-cd_k}.

We first record two bounds. Recall from Lemma 5.14 that the entries of Q are multipliers of degree one and that the soul of d_k is bounded by Cv. Using also $(d_k)_{\emptyset} \geq v - 1 - k$, from (5.10), we obtain

$$|e^{-cd_k}|_{1;1} \leq e^{-c(d_k)_{\emptyset}} \sum_{j \leq 2k^2} \frac{(Ccv)^j}{j!} \leq C(1 + cv)^{2k^2} e^{-c(v-1-k)}, \tag{5.29}$$

where we used that a product of more than $2k^2$ souls of even elements vanishes. Moreover, in the translated charts, the coefficients of $\mathfrak{q}$ are bounded by $Cv^{1/2}$, by the construction in the proof of Lemma 5.4, since $\|g^{-1}q_{\mathrm{A}}g\|_{\mathrm{op}} \le Cv^{1/2}$ for the choice of g in the proof of Lemma 5.14.

Step 1. We show that I is constant. Let ψ_R be the compactly supported cutoff (5.21), and set $w = \mathfrak{o}Ae^{-cd_k}$. By Lemma 5.18, Lemma 3.4(b), and $\mathfrak{q}A = 0$, we have $\mathfrak{q}e^{-cd_k} = 0$ and $\mathfrak{q}w = d_kAe^{-cd_k}$, as in the proof of Lemma 3.18. The invariance of the integral and the Leibniz rule give

$$\int \psi_R\, d_kAe^{-cd_k} = \int \psi_R\, \mathfrak{q}w = -\int (\mathfrak{q}\psi_R)\, w.$$

The cutoff ψ_R depends only on the body point. Using Lemma 5.1(b), we find that $|\mathfrak{q}\psi_R|_{1;1} \le CvR^{-1}$ and that $\mathfrak{q}\psi_R$ is supported where $v \ge e^R$. The superfunctions vw and $d_kAe^{-cd_k}$ are integrable, locally uniformly in $c > 0$, by (5.29) and $|\mathfrak{o}|_{1;1} \le Cv$. Dominated convergence shows that I is differentiable, and letting $R \to \infty$ in the last identity, we obtain

$$I'(c) = -\int d_kAe^{-cd_k} = -\lim_{R\to\infty} \int \psi_R\, d_kAe^{-cd_k} = 0.$$

Step 2. We consider the limit $c \downarrow 0$. For $0 < c \le 1$, the bound (5.29) implies that $|e^{-cd_k}|_{1;1} \le Cv^{2k^2}$, and the coefficients of e^{-cd_k} converge to those of 1 as $c \downarrow 0$. Since A has polynomially integrable coefficients, dominated convergence yields

$$\lim_{c\downarrow 0} I(c) = \int A.$$

Step 3. We consider the limit $c \to \infty$. We use the exponential chart at Λ. Let $\mathcal{O}$ be a small neighborhood of zero in $\mathfrak{m}$, and write $Y = Y_0 + \Theta' \in \mathcal{O}$, with Y_0 in the real subspace of $\mathfrak{m}_0$ described in Section 5.2 and $\Theta' \in \mathfrak{m}_1$ with coefficients in the odd generators. The superfunctions

$$Q = e^{Y}\Lambda e^{-Y} = \Lambda e^{-2Y}$$

define coordinates on a neighborhood of Λ, by Lemma 3.14, since at $Y = 0$ this change of coordinates agrees with the regular coordinates to first order. In these coordinates,

$$d_k = \frac{1}{2}\,\mathrm{STr}(e^{-2Y}) = \sum_{n\ge 1} \frac{2^{2n-1}}{(2n)!}\,\mathrm{STr}(Y^{2n}),$$

since odd powers of $Y \in \mathfrak{m}$ have vanishing flavor blocks RR and AA, and therefore vanishing supertrace. By Berezin's formula (Theorem 3.15), the invariant integral has a density $\rho_{\mathcal{O}}(Y)$ in these coordinates, and its body at $Y = 0$ is a nonzero constant $\rho_{\mathcal{O}}(0)$, since ρ has nonzero body.

Since $(d_k)_\varnothing = \mathrm{STr}(Y_0^2) + O(|Y_0|^4)$, where $|Y_0|^2 = \mathrm{STr}(Y_0^2)$, we may choose $\mathcal{O}$ so small that $(d_k)_\varnothing \geq a|Y_0|^2$ on $\mathcal{O}$ for some $a > 0$. Let ϕ be a smooth function on the underlying manifold with compact support in $\mathcal{O}$ and equal to one near zero, and split

$$I(c) = \int \phi A e^{-cd_k} + \int (1-\phi) A e^{-cd_k}.$$

On the support of $1-\phi$, the body of d_k is at least $\max(\delta, v-1-k)$ for some $\delta > 0$, and (5.29) and dominated convergence show that the second integral tends to zero.

In the first integral, we change variables by $Y = c^{-1/2}Y'$ in all even and odd coordinates. Since $\mathfrak{m}$ has as many even as odd dimensions, this change of variables has Berezinian one (Lemma 3.6(c)), and

$$\int \phi A e^{-cd_k} = \int_{\mathfrak{m}} \big(\phi \rho_{\mathcal{O}} A\big)(c^{-1/2}Y')\, e^{-\mathrm{STr}(Y'^2) - r_c(Y')}\, \mathrm{d}Y',$$
$$r_c(Y') = \sum_{n \geq 2} \frac{2^{2n-1}}{(2n)!}\, c^{1-n}\, \mathrm{STr}(Y'^{2n}).$$

As $c \to \infty$, every coefficient of the integrand converges to the corresponding coefficient of $A_\varnothing(\Lambda)\rho_{\mathcal{O}}(0)e^{-\mathrm{STr}(Y'^2)}$. Indeed, the terms of positive degree in the odd coordinates of A, $\rho_{\mathcal{O}}$, and ϕ come with positive powers of $c^{-1/2}$, and the bodies converge to their values at zero. On the support of $\phi(c^{-1/2}\,\cdot)$, the body of the exponent $\mathrm{STr}(Y'^2) + r_c(Y') = c\, d_k(c^{-1/2}Y')$ is at least $a|Y_0'|^2$, and the souls of this exponent and of $(\phi\rho_{\mathcal{O}}A)(c^{-1/2}Y')$ are bounded uniformly in $c \geq 1$. We obtain the integrable majorant $Ce^{-a|Y_0'|^2}$ for the coefficients, and dominated convergence yields

$$\lim_{c\to\infty} I(c) = A_\varnothing(\Lambda)\, \gamma_k, \qquad \gamma_k = \rho_{\mathcal{O}}(0) \int_{\mathfrak{m}} e^{-\mathrm{STr}(Y'^2)}\, \mathrm{d}Y'.$$

The Gaussian integral is nonzero, by Proposition 3.8, since $\mathrm{STr}(Y'^2) = 2\,\mathrm{STr}(Y'_{\mathrm{RA}}Y'_{\mathrm{AR}})$ is nondegenerate in the odd coordinates.

We now conclude. Applying Steps 1 and 3 to $A = 1$, which is possible since they use only the integrability of e^{-cd_k} and $d_k e^{-cd_k}$, we find that $\int e^{-cd_k} = \gamma_k$ for all $c > 0$. In particular, the normalization (5.12) is possible, since $\gamma_k \neq 0$, and it implies that $\gamma_k = 1$, so that

$$\int e^{-cd_k} = 1, \qquad c > 0. \tag{5.30}$$

For general A, the three steps give

$$\int A = \lim_{c\downarrow 0} I(c) = \lim_{c\to\infty} I(c) = A_\varnothing(\Lambda).$$

The statement for matrices follows by applying the scalar statement to each entry. □

We now prove that the partition function is one at the base point, and that the semigroup preserves the symmetries of the constant function. The point is that the generator commutes with the vector fields of the stabilizer $\mathfrak{k}$.

Corollary 5.20. *Let $X > 0$ and $z > 0$. Then*

$$\mathcal{B}_k\big(e^{-a\mathsf{G}_{X,z}}1, e^{-(T-a)\mathsf{G}_{X,z}}1\big) = \int e^{-T\mathsf{G}_{X,z}}1 = 1, \qquad 0 < a < T.$$

In particular, $\mathcal{Z}^{(k)}_{X,\varepsilon}(0, \varepsilon\mathbf{1}) = 1$ for all $\varepsilon > 0$. Moreover, $e^{-t\mathsf{G}_{X,z}}1$ is invariant under $\mathfrak{k}$ for all $t \geq 0$.

Proof. Write $\mathsf{G} = \mathsf{G}_{X,z}$, and let $Y \in \mathfrak{k}$ be homogeneous. Since d_k is invariant under $\mathfrak{k}$, Lemma 5.8(a) implies that $\widehat{Y}\mathsf{G}f = \mathsf{G}\widehat{Y}f$ for all smooth f. Fix $c > 0$. The superfunctions $\mathsf{G}^n e^{-cd_k}$, $n \geq 0$, are rapidly decreasing, by (5.29) and the same bound for the derivatives. Applying (5.25) inductively, we find that $e^{-cd_k} \in D(\mathsf{G}^n_{(p)})$ for all p and n. Let

$$\Phi(t) = e^{-t\mathsf{G}}e^{-cd_k}, \qquad \Phi_1(t) = \widehat{Y}\Phi(t).$$

We first show that $\Phi_1 = 0$. The construction of Lemma 5.4 shows that the operator $\widehat{Y}$ is of first order, with coefficients bounded by Cv in the translated charts. In particular, $\|\widehat{Y}f\|_{p-1} \leq C\|f\|_{V_p}$. The norm $\|f\|_{V_p}$ is bounded by the graph norm of $\mathsf{G}_{(p)}$, by (C.4). Since $\Phi(0) \in D(\mathsf{G}^2_{(p)})$, the map $t \mapsto \Phi(t)$ is continuously differentiable on $[0, \infty)$ into $D(\mathsf{G}_{(p)})$, with derivative $-\mathsf{G}\Phi(t)$. It follows that Φ_1 is continuously differentiable into L^2_{p-1}, and

$$\Phi_1'(t) = -\widehat{Y}\mathsf{G}\Phi(t) = -\mathsf{G}\Phi_1(t),$$

where we used that $\Phi(t)$ is smooth.

The superfunction $\Phi(t)$ is rapidly decreasing for $t > 0$, by Proposition 5.15(b), hence so is $\Phi_1(t)$, and $\Phi_1(t) \in D(\mathsf{G}_{(p-1)})$ by (5.25). Moreover $\Phi_1(0) = \widehat{Y}e^{-cd_k} = 0$, since d_k is invariant under $\mathfrak{k}$. Thus Φ_1 is a strong solution of $\Phi_1' = -\mathsf{G}_{(p-1)}\Phi_1$ with zero initial value, and for $0 < s < t$,

$$\frac{\mathrm{d}}{\mathrm{d}s}e^{-(t-s)\mathsf{G}}\Phi_1(s) = e^{-(t-s)\mathsf{G}}\big(\mathsf{G}\Phi_1(s) + \Phi_1'(s)\big) = 0. \tag{5.31}$$

Integrating over $s \in (0, t)$, we obtain $\Phi_1(t) = 0$.

Next, we show that $\int \Phi(t)$ is independent of t. Since $\Phi(t)$ is rapidly decreasing and $-\mathsf{G}\Phi(t)$ is its derivative in every L^2_p, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \Phi(t) = -X\int \mathsf{C}_k\Phi(t) - z\int d_k\Phi(t).$$

The first integral on the right vanishes. Indeed, Lemma 5.8(b) gives

$$\int \big(\mathsf{C}_k\Phi(t)\big)\psi_R = \int \Phi(t)\,\mathsf{C}_k\psi_R, \qquad |\mathsf{C}_k\psi_R|_{1;1} \leq CR^{-1},$$

where the bound holds since the coefficients of C_k are bounded in translated charts and the derivatives of ψ_R of order one and two are bounded by CR^{-1}. The right side tends to zero as $R \to \infty$. For the second integral, we note that $d_k\Phi(t)$ is $\mathfrak{q}_\mathrm{A}$-closed, by the first step with $Y = q_\mathrm{A}$ and Lemma 5.18, and that it has polynomially integrable coefficients. Since $(d_k)_\varnothing(\Lambda) = 0$, Lemma 5.19 gives

$$\int d_k\Phi(t) = \big(d_k\Phi(t)\big)_\varnothing(\Lambda) = 0.$$

Together with (5.30), we conclude that

$$\int \Phi(t) = \int e^{-cd_k} = 1, \qquad t \geq 0.$$

Finally, we let $c \downarrow 0$. Let $q > q_1/2$. Using (5.29) and dominated convergence, we find that $e^{-cd_k} \to 1$ in L^2_{-q}. Using Proposition 5.15(d), (c), and the boundedness of $\mathcal{B}_k$ on $L^2_q \times L^2_{-q}$, we obtain

$$\mathcal{B}_k\big(e^{-a\mathsf{G}}1, e^{-(T-a)\mathsf{G}}1\big) = \mathcal{B}_k(e^{-T\mathsf{G}}1, 1) = \lim_{c\downarrow 0} \mathcal{B}_k(e^{-T\mathsf{G}}1, e^{-cd_k}) = \lim_{c\downarrow 0} \int e^{-T\mathsf{G}}e^{-cd_k} = 1.$$

The case $z = 2\pi\varepsilon$, $T = 1$ is the identity $\mathcal{Z}^{(k)}_{\mathrm{X},\varepsilon}(0, \varepsilon\mathbf{1}) = 1$, since $\mathsf{G}_{X,\Gamma_0} = \mathsf{G}_{X,2\pi\varepsilon}$.

For the last assertion, let $t > 0$. As $c \downarrow 0$, the superfunctions $e^{-t\mathsf{G}}e^{-cd_k} = e^{-t\mathsf{G}/2}\big(e^{-t\mathsf{G}/2}e^{-cd_k}\big)$ converge to $e^{-t\mathsf{G}}1$ in the graph norm of $\mathsf{G}_{(-q)}$, since $e^{-t\mathsf{G}/2}$ maps L^2_{-q} boundedly into $D(\mathsf{G}_{(-q)})$. Since $\widehat{Y}e^{-t\mathsf{G}}e^{-cd_k} = 0$ and $\|\widehat{Y}f\|_{-q-1} \leq C\|f\|_{V_{-q}}$, we obtain $\widehat{Y}e^{-t\mathsf{G}}1 = 0$ in the limit. The case $t = 0$ is trivial. □

We will also use localization for matrices of superfunctions. For a matrix $F = (F_{ij})$ of superfunctions indexed by the retarded indices, we say that F is *retarded-covariant* if $\vartheta_Y\widehat{Y}F = [\vartheta_Y Y_{\mathrm{RR}}, F]$ for all homogeneous $Y \in \mathfrak{k}$, where $\widehat{Y}$ acts entrywise and Y_{RR} is the retarded flavor block of Y, and *advanced-invariant* if $\widehat{Y}F = 0$ for all $Y \in \mathfrak{k}_\mathrm{A}$. For instance, Q_{RR} is retarded-covariant and advanced-invariant, by (5.6), and so is $f\mathrm{Id}_\mathrm{R}$ for every superfunction f invariant under $\mathfrak{k}$. Advanced-covariant and retarded-invariant matrices are defined in the same way, with the roles of the flavors exchanged.

Corollary 5.21. *Let F be a retarded-covariant and advanced-invariant matrix of smooth superfunctions. Then there exists $c_\Lambda \in \mathbb{C}$ such that $F_\varnothing(\Lambda) = c_\Lambda \mathrm{Id}_\mathrm{R}$. If moreover f is a smooth superfunction invariant under $\mathfrak{k}$ such that fF has polynomially integrable coefficients, then*

$$\int fF = f_\varnothing(\Lambda)\, c_\Lambda\, \mathrm{Id}_\mathrm{R}.$$

The same holds with the roles of the retarded and advanced flavors exchanged.

Proof. The entries of fF are invariant under $\mathfrak{k}_A$, hence $\mathfrak{q}_A$-closed by Lemma 5.18, and Lemma 5.19 implies that

$$\int fF = f_\varnothing(\Lambda) F_\varnothing(\Lambda).$$

It remains to identify $F_\varnothing(\Lambda)$.

Let $Y \in \mathfrak{k}_0$ be supported on the retarded flavors. Since Λ is fixed by e^{sY}, the body of $\widehat{Y}F$ vanishes at Λ, and covariance implies that $[Y_{RR}, F_\varnothing(\Lambda)] = 0$ for all such Y. In other words, $F_\varnothing(\Lambda)$ commutes with all block diagonal matrices of the retarded flavors. Its off-diagonal blocks have odd entries and vanishing body. It follows that $F_\varnothing(\Lambda) = \mathrm{diag}(c_B \mathrm{Id}_k | c_F \mathrm{Id}_k)$ for some $c_B, c_F \in \mathbb{C}$. Now let $Y \in \mathfrak{k}_1$ be the supermatrix with a single entry one in the block Y^{BF}_{RR}. At $g = \mathrm{Id}$, the matrices A and Θ_1 in the construction of Lemma 5.4 have degree at least one in the χ_a, and the body of $\widehat{Y}F$ vanishes at Λ as well. Taking the body at Λ of the BF block of the covariance relation, we find $\vartheta_Y(c_F - c_B) Y^{BF}_{RR} = 0$, and we conclude that $c_B = c_F$. □

SECTION 6

The sigma model of rank one

In this section, we treat the sigma model of rank one, in which the objects of Section 5 can be computed explicitly. We show that the semigroup acts on invariant superfunctions as the semigroup of the singlet diffusion, so that its value on the constant function is the singlet state $\Psi_t^{X,z}$, and we prove the properties of the kernels K_X, $\widetilde{K}_X$, and I_X that Definition 1.4 uses. We then develop the calculus of rank-one pairings, which evaluates the pairings of singlet states with entries of Q inserted, and we show how K_X and I_X arise from the derivatives of the partition function in its source, and $\widetilde{K}_X$ from an insertion at a single site. Part III uses these results throughout, and Part IV does not use them. From Section 5, we use its definitions at $k = 1$, the symmetries and the Casimir operator of $\mathcal{M}_1$ through Lemmas 5.4, 5.8 and 5.11, the bounds of Lemmas 5.1 and 5.14, the semigroups of Proposition 5.15 with the regularity of semigroup products in Lemma 5.17, and the localization identity Lemma 5.19, together with the supercharges of Lemma 5.18 and its consequences in Corollaries 5.20 and 5.21. Lemma 6.14, by which the partition function extends to an entire function of the sources, holds at every rank, and Section 8.4 uses it in this generality.

Throughout this section, $k = 1$, and we order the indices as $(\mathrm{RB}, \mathrm{RF}, \mathrm{AB}, \mathrm{AF})$. Recall that the underlying manifold of $\mathcal{M}_1$ is the product $H_1 \times S_1$ of the hyperbolic plane and the sphere, and that $\mathcal{M}_1$ has four odd coordinates. We introduce polar coordinates on $\mathcal{M}_1$, in which the invariant integral and the generator $\mathsf{G}_{X,z}$ reduce to explicit expressions in the radial coordinates $(B, F) \in D$. The polar coordinates are singular at the point Λ, which corresponds to the corner $(1, 1)$ of D, and the corresponding corner term, a boundary term of the kind described in Section 3.6, is computed by localization.

We write $p = (B, F)$ for points of D, as in Definition 1.4. We set $d = B - F$, and $D^\circ = (1, \infty) \times (-1, 1)$. As in Section 1.2, we call $(1, 1)$ the corner of D. We recall from Definition 1.4 the operator Δ_D and the diffusion $(\tilde{B}_t, \tilde{F}_t)$ with generator $X\Delta_D$, which we call the *radial diffusion*, and the semigroup $e^{t(X\Delta_D - zd)}$, and we define the *singlet operator*

$$\Delta_D^{\mathrm{s}} = d\,\Delta_D\,d^{-1} = \Delta_D - \frac{2(B^2-1)}{d}\partial_B + \frac{2(1-F^2)}{d}\partial_F \tag{6.1}$$

on D minus the corner. The second form follows from $\Delta_D d^{-1} = 0$, which one checks directly using $\partial_B d = 1$, $\partial_F d = -1$, and $(B^2 - 1) + (1 - F^2) = d(B + F)$. We will use the identity $\Delta_D^{\mathrm{s}}(dh) = d\,\Delta_D h$ repeatedly.

6.1 Polar coordinates

For $p \in D$ and $\theta_{\mathrm{B}}, \theta_{\mathrm{F}} \in \mathbb{R}/2\pi\mathbb{Z}$, we let $Q_0 = Q_0(B, F, \theta_{\mathrm{B}}, \theta_{\mathrm{F}})$ denote the block diagonal supermatrix with blocks

$$Q_0^{\mathrm{BB}} = \begin{pmatrix} B & \sqrt{B^2-1}\, e^{\mathrm{i}\theta_{\mathrm{B}}} \\ -\sqrt{B^2-1}\, e^{-\mathrm{i}\theta_{\mathrm{B}}} & -B \end{pmatrix}, \qquad Q_0^{\mathrm{FF}} = \begin{pmatrix} F & \mathrm{i}\sqrt{1-F^2}\, e^{\mathrm{i}\theta_{\mathrm{F}}} \\ -\mathrm{i}\sqrt{1-F^2}\, e^{-\mathrm{i}\theta_{\mathrm{F}}} & -F \end{pmatrix},$$

which is the point of $\mathcal{M}_1^0$ with coordinates $(B, \theta_{\mathrm{B}}, F, \theta_{\mathrm{F}})$ of Section 5.2. Let $\xi_1, \dots, \xi_4$ be odd generators, and let Θ_ξ denote the odd supermatrix in $\mathfrak{k}_1$ with entries ξ_1 at $(\mathrm{RB}, \mathrm{RF})$, ξ_2 at $(\mathrm{RF}, \mathrm{RB})$, ξ_3 at $(\mathrm{AB}, \mathrm{AF})$, and ξ_4 at $(\mathrm{AF}, \mathrm{AB})$. We define the *polar coordinates* by

$$Q = e^{-\Theta_\xi} Q_0(B, F, \theta_{\mathrm{B}}, \theta_{\mathrm{F}})\, e^{\Theta_\xi}, \tag{6.2}$$

and we set $n_{\mathrm{R}} = \xi_1\xi_2$ and $n_{\mathrm{A}} = \xi_3\xi_4$. These are the coordinates of Efetov [39], written as in [95, Section 2.4], up to a linear change of the odd variables.

Lemma 6.1. *(a) The map* (6.2) *identifies the superfunctions on $\mathcal{M}_1$, restricted to the open set of body points with $B > 1$ and $|F| < 1$, with the superfunctions of $(B, F, \theta_{\mathrm{B}}, \theta_{\mathrm{F}})$ in $D^\circ \times (\mathbb{R}/2\pi\mathbb{Z})^2$ and of $\xi_1, \dots, \xi_4$. The same holds near the edges $\{B = 1\}$ and $\{F = \pm 1\}$ of D minus the corner, after the polar coordinates (B, θ_{B}), respectively (F, θ_{F}), are replaced by Cartesian coordinates of the plane, respectively of the sphere.*

(b) In polar coordinates, we have $Q^2 = \mathrm{Id}$ and

$$\begin{aligned} Q_{\mathrm{RR}}^{\mathrm{BB}} &= B + d\, n_{\mathrm{R}}, & Q_{\mathrm{RR}}^{\mathrm{FF}} &= F + d\, n_{\mathrm{R}}, & Q_{\mathrm{RR}}^{\mathrm{BF}} &= d\,\xi_1, & Q_{\mathrm{RR}}^{\mathrm{FB}} &= -d\,\xi_2, \\ Q_{\mathrm{AA}}^{\mathrm{BB}} &= -B - d\, n_{\mathrm{A}}, & Q_{\mathrm{AA}}^{\mathrm{FF}} &= -F - d\, n_{\mathrm{A}}, & Q_{\mathrm{AA}}^{\mathrm{BF}} &= -d\,\xi_3, & Q_{\mathrm{AA}}^{\mathrm{FB}} &= d\,\xi_4. \end{aligned}$$

In particular, $d_1 = \frac{1}{2}\,\mathrm{STr}(\Lambda Q) = d$ and $\frac{1}{4}\,\mathrm{STr}((\Lambda Q)^2) = B^2 - F^2$, with no odd terms.

(c) Let $[f]_{\mathrm{top}}$ denote the coefficient of $\xi_1\xi_2\xi_3\xi_4$ in f, and $\langle f \rangle_{\mathrm{top}}$ its average over θ_{B} and θ_{F} with respect to $\frac{\mathrm{d}\theta_{\mathrm{B}}}{2\pi}\frac{\mathrm{d}\theta_{\mathrm{F}}}{2\pi}$. Then

$$\begin{aligned} [Q_{\mathrm{RR}}^{\mathrm{BB}} Q_{\mathrm{AA}}^{\mathrm{BB}}]_{\mathrm{top}} &= [Q_{\mathrm{RR}}^{\mathrm{FF}} Q_{\mathrm{AA}}^{\mathrm{FF}}]_{\mathrm{top}} = -d^2, \\ \langle Q_{\mathrm{RA}}^{\mathrm{BB}} Q_{\mathrm{AR}}^{\mathrm{BB}} \rangle_{\mathrm{top}} &= \langle Q_{\mathrm{AR}}^{\mathrm{FB}} Q_{\mathrm{RA}}^{\mathrm{BF}} \rangle_{\mathrm{top}} = -(B^2 - F^2). \end{aligned}$$

Proof. We begin with (b). The flavor blocks of Θ_ξ are $(\Theta_\xi)_{\mathrm{RR}} = \left(\begin{smallmatrix} 0 & \xi_1 \\ \xi_2 & 0 \end{smallmatrix}\right)$ and $(\Theta_\xi)_{\mathrm{AA}} = \left(\begin{smallmatrix} 0 & \xi_3 \\ \xi_4 & 0 \end{smallmatrix}\right)$, and since $(\xi_1\xi_2)^2 = 0$,

$$e_{\mathrm{RR}}^{\pm\Theta_\xi} = \begin{pmatrix} 1 + \frac{1}{2}\xi_1\xi_2 & \pm\xi_1 \\ \pm\xi_2 & 1 - \frac{1}{2}\xi_1\xi_2 \end{pmatrix},$$

and similarly for AA. Since Θ_ξ is block diagonal in the flavors, we have

$$Q_{\mathrm{RR}} = e_{\mathrm{RR}}^{-\Theta_\xi}\, \mathrm{diag}(B, F)\, e_{\mathrm{RR}}^{\Theta_\xi},$$

and multiplying out gives the first line; the second line is the same computation with $\mathrm{diag}(-B,-F)$. For the supertraces, we use $\mathrm{STr}(\Lambda Q) = \mathrm{STr}\, Q_{\mathrm{RR}} - \mathrm{STr}\, Q_{\mathrm{AA}}$, the identity

$$\mathrm{STr}((\Lambda Q)^2) = 2\,\mathrm{STr}(Q_{\mathrm{RR}}^2) + 2\,\mathrm{STr}(Q_{\mathrm{AA}}^2) - 2\,\mathrm{STr}(\mathrm{Id}),$$

which follows from $Q^2 = \mathrm{Id}$, and $\mathrm{STr}(Q_{\mathrm{RR}}^2) = \mathrm{STr}(\mathrm{diag}(B,F)^2)$, which follows from cyclicity. Part (c) is obtained by multiplying out the products in the Grassmann algebra with four generators; for the first identity, only the product of the terms $d\,n_{\mathrm{R}}$ and $-d\,n_{\mathrm{A}}$ contributes. (We have also checked all these identities by computer algebra.)

We now prove (a). The body map is the product of the polar coordinates of the plane and of the sphere, which is a diffeomorphism of $D^\circ \times (\mathbb{R}/2\pi\mathbb{Z})^2$ onto the set $\mathcal{O}^\circ$ of body points with $B > 1$ and $|F| < 1$. Let s be a section over an open set $V \subset \mathcal{O}^\circ$, with coordinates (x,χ). By (6.2), the regular coordinates are superfunctions of the polar coordinates over the preimage of V, namely $x = y(B,F,\theta_{\mathrm{B}},\theta_{\mathrm{F}};\xi)$ and $\chi = \eta(B,F,\theta_{\mathrm{B}},\theta_{\mathrm{F}};\xi)$, with y even and η odd. We apply Lemma 3.14 with $m = r = 4$, with $(B,F,\theta_{\mathrm{B}},\theta_{\mathrm{F}};\xi_1,\dots,\xi_4)$ as the old coordinates, and with (y,η) as the new ones. The body map $y_\varnothing$ is a diffeomorphism, and it remains to show that the body of the matrix $(\partial_{\xi_a}\eta_b)$ is invertible, or equivalently that the linear map from $(\xi_1,\dots,\xi_4)$ to the terms of degree one of the regular odd coordinates Θ is injective. Lemma 3.14 then shows that $f \mapsto f(y,\eta)$ identifies the superfunctions over V with the superfunctions of the polar coordinates, which is (a).

Recall from Remark 5.3 that the part of degree one of the odd entries of Q is the matrix $-2g\Lambda\Theta g^{-1}$, which is linear in Θ. Suppose that the terms of degree one of Θ vanish. Then so do those of $Q_{\mathrm{RR}}^{\mathrm{BF}} = d\xi_1$, $Q_{\mathrm{RR}}^{\mathrm{FB}} = -d\xi_2$, $Q_{\mathrm{AA}}^{\mathrm{BF}} = -d\xi_3$, and $Q_{\mathrm{AA}}^{\mathrm{FB}} = d\xi_4$, and $d > 0$ forces $\xi = 0$. The same argument applies near the edges with Cartesian coordinates, since $d > 0$ there except at the corner. □

We first identify the invariant integral away from the corner. Let $\mathcal{O}$ denote the set of body points with $d > 0$, that is, $\mathcal{M}_1^0$ minus the point Λ, which is connected, and let $\mathcal{O}^\circ \subset \mathcal{O}$ denote the set of body points with $B > 1$ and $|F| < 1$, on which the polar coordinates are defined. The following lemma says that, away from the corner, the invariant integral has the density d^{-2} in polar coordinates, up to a constant factor.

Lemma 6.2. *There exists a constant c_1 such that*

$$\int f = c_1 \int_D \frac{\langle f\rangle_{\mathrm{top}}}{d^2}\,\mathrm{d}B\,\mathrm{d}F \tag{6.3}$$

for all smooth superfunctions f on $\mathcal{M}_1$ whose coefficients have compact support in $\mathcal{O}$.

Proof. For f as in the statement, we set

$$J(f) = \int_{D^\circ\times(\mathbb{R}/2\pi\mathbb{Z})^2} \frac{[f]_{\mathrm{top}}}{d^2}\,\mathrm{d}B\,\mathrm{d}F\,\frac{\mathrm{d}\theta_{\mathrm{B}}}{2\pi}\,\frac{\mathrm{d}\theta_{\mathrm{F}}}{2\pi} = \int_D \frac{\langle f\rangle_{\mathrm{top}}}{d^2}\,\mathrm{d}B\,\mathrm{d}F,$$

where $[f]_{\mathrm{top}}$ is computed in polar coordinates. Near the edges of D, we use the Cartesian coordinates of Lemma 6.1(a). Since $\mathrm{d}B\,\mathrm{d}\theta_{\mathrm{B}} = B^{-1}\,\mathrm{d}^2w$ with $w = \sqrt{B^2-1}\,e^{\mathrm{i}\theta_{\mathrm{B}}}$, and the analogous factor for the sphere is smooth, J is a Berezin integral with a smooth and nonvanishing density in each of these coordinate systems over $\mathcal{O}$.

By Lemma 6.1(a) and Theorem 3.15, the same holds for the invariant integral, and hence there is an even superfunction σ on $\mathcal{O}$ with nonvanishing body such that

$$\int f = J(\sigma f)$$

for all f with compact support in $\mathcal{O}$. We show that σ is constant.

We first show that J is invariant under $\mathfrak{k}$, that is, $J(\widehat{Y}f) = 0$ for all homogeneous $Y \in \mathfrak{k}$ and all f with compact support in $\mathcal{O}^\circ$. By the uniqueness in Lemma 5.4(a), the even derivation $\vartheta_Y\widehat{Y}$ maps f to $f(e^{\vartheta_Y Y}Qe^{-\vartheta_Y Y}) - f(Q)$, to first order in ϑ_Y, since both derivations map Q to $[\vartheta_Y Y, Q]$. At rank one, K_0 consists of diagonal unitary matrices. Conjugation by an element of K_0 commutes with Λ, shifts the phases θ_{B} and θ_{F} by constants, and multiplies each ξ_a by a phase, with inverse phases for ξ_1 and ξ_2 and for ξ_3 and ξ_4. It therefore preserves $\mathrm{d}\theta_{\mathrm{B}}\,\mathrm{d}\theta_{\mathrm{F}}$ and the coefficient of $\xi_1\xi_2\xi_3\xi_4$, and J is invariant under K_0. Differentiating, we obtain $J(\widehat{Y}f) = 0$ for Y in the Lie algebra of K_0, and for all $Y \in \mathfrak{k}_0$ by complex linearity. For the odd generator $Y = E_{\mathrm{RB,RF}}$ of $\mathfrak{k}_1$ and $\vartheta = \vartheta_Y$, a direct computation with the matrices $e^{\pm\Theta_\xi}_{\mathrm{RR}}$ of the proof of Lemma 6.1 gives

$$e^{\vartheta Y}e^{-\Theta_\xi} = e^{-\Theta_{\xi'}}\Big(\big(1 - \tfrac{1}{2}\vartheta\xi_2\big)\mathrm{Id}_{\mathrm{R}} + \mathrm{Id}_{\mathrm{A}}\Big), \qquad \xi' = (\xi_1 - \vartheta, \xi_2, \xi_3, \xi_4),$$

to first order in ϑ. Conjugation by the last factor multiplies the flavor block RA of Q_0 by $1 - \frac{1}{2}\vartheta\xi_2$ and the block AR by its inverse, which is the shift of both phases by $\frac{\mathrm{i}}{2}\vartheta\xi_2$. In polar coordinates, we conclude that

$$\vartheta\widehat{Y} = -\vartheta\,\partial_{\xi_1} + \tfrac{\mathrm{i}}{2}\vartheta\xi_2\big(\partial_{\theta_{\mathrm{B}}} + \partial_{\theta_{\mathrm{F}}}\big).$$

Neither coefficient depends on the variable that it differentiates, so $J(\vartheta\widehat{Y}f) = 0$, by Berezin integration by parts in ξ_1 (Lemma 3.6(a)) and by periodicity in the phases. The other three generators $E_{\mathrm{RF,RB}}$, $E_{\mathrm{AB,AF}}$, and $E_{\mathrm{AF,AB}}$ of $\mathfrak{k}_1$ are treated in the same way, with ξ_2, ξ_3, or ξ_4 shifted.

Next, we show that σ is a function of (B, F). Since σ is even, the invariance of both integrals gives, for $Y \in \mathfrak{k}$ and f with compact support in $\mathcal{O}^\circ$,

$$J\big((\widehat{Y}\sigma)f\big) = J\big(\widehat{Y}(\sigma f)\big) - J(\sigma\,\widehat{Y}f) = 0 - \int \widehat{Y}f = 0,$$

and therefore $\widehat{Y}\sigma = 0$ for all $Y \in \mathfrak{k}$. Write $\Theta_\xi = \sum_a \xi_a Y_a$ with $Y_a \in \mathfrak{k}_1$. By (5.6) and the chain rule, we have $\frac{\mathrm{d}}{\mathrm{d}s}\sigma(e^{-s\Theta_\xi}Q_0e^{s\Theta_\xi}) = -\sum_a \xi_a(\widehat{Y}_a\sigma)(e^{-s\Theta_\xi}Q_0e^{s\Theta_\xi}) = 0$. Hence σ has no

odd terms in polar coordinates, and by the invariance under K_0, it does not depend on the phases. Thus $\sigma = \sigma(B, F)$ on $\mathcal{O}^\circ$.

We now show that σ is constant, with the invariance of the integral under the even generator $Y_0 = E_{\mathrm{RB,AB}}$ of $\mathfrak{m}_0$. By Lemma 6.1(b), the superfunctions d_1 and $\frac{1}{4}\operatorname{STr}((\Lambda Q)^2)$ equal d and $m = B^2 - F^2$ in polar coordinates, without odd terms. Since $\partial(d, m)/\partial(B, F) = 2d \neq 0$ and $(d, m/d)$ determines (B, F), the pair (d, m) is a system of coordinates on D°, whose coordinate vector fields are

$$\partial_d = -\frac{F\partial_B + B\partial_F}{d}, \qquad \partial_m = \frac{\partial_B + \partial_F}{2d}.$$

In particular, every $\phi \in C_c^\infty(D^\circ)$ is a smooth function of (d, m). Substituting d_1 and $\frac{1}{4}\operatorname{STr}((\Lambda Q)^2)$ for d and m, we obtain a superfunction on $\mathcal{O}^\circ$ that equals $\phi(B, F)$ in polar coordinates, with no odd terms, and we denote it also by ϕ.

We now compute the action of $\widehat{Y}_0$. By Lemma 5.4(b), we have $\widehat{Y}_0 Q = [Y_0, Q]$. Using $\operatorname{STr}(MY_0) = M_{\mathrm{AB,RB}}$ for supermatrices M and the cyclicity of the supertrace, we find

$$\begin{gathered} \widehat{Y}_0 d_1 = Q_{\mathrm{AB,RB}}, \qquad \widehat{Y}_0 \tfrac{1}{4}\operatorname{STr}\big((\Lambda Q)^2\big) = \tfrac{1}{2}[Q, \Lambda Q \Lambda]_{\mathrm{AB,RB}}, \\ \widehat{Y}_0 Q_{\mathrm{RB,AB}} = Q^{\mathrm{BB}}_{\mathrm{AA}} - Q^{\mathrm{BB}}_{\mathrm{RR}}, \qquad \widehat{Y}_0 Q^{\mathrm{FF}}_{\mathrm{RR}} = 0. \end{gathered} \tag{6.4}$$

Next, we compute two phase averages. For a superfunction A on $\mathcal{O}^\circ$, let $\langle A \rangle$ denote the average of its coefficients over the phases. Since Λ commutes with Θ_ξ, we have $[Q, \Lambda Q \Lambda] = e^{-\Theta_\xi}[Q_0, \Lambda Q_0 \Lambda] e^{\Theta_\xi}$. Multiplying the 2×2 blocks, we obtain $[Q_0, \Lambda Q_0 \Lambda]_{\mathrm{AR}} = 4\operatorname{diag}(B, F)(Q_0)_{\mathrm{AR}}$. We write $Q_{\mathrm{RA}} = e^{-\Theta_\xi}_{\mathrm{RR}}(Q_0)_{\mathrm{RA}} e^{\Theta_\xi}_{\mathrm{AA}}$ and $Q_{\mathrm{AR}} = e^{-\Theta_\xi}_{\mathrm{AA}}(Q_0)_{\mathrm{AR}} e^{\Theta_\xi}_{\mathrm{RR}}$, and we multiply out. The terms with the phases $e^{\pm \mathrm{i}(\theta_{\mathrm{B}} - \theta_{\mathrm{F}})}$ average to zero, and we find

$$\begin{aligned} \big\langle Q_{\mathrm{AB,RB}} Q_{\mathrm{RB,AB}} \big\rangle &= -(B^2 - 1)(1 + n_{\mathrm{R}})(1 + n_{\mathrm{A}}) - (1 - F^2)\, n_{\mathrm{R}} n_{\mathrm{A}}, \\ \big\langle [Q, \Lambda Q \Lambda]_{\mathrm{AB,RB}} Q_{\mathrm{RB,AB}} \big\rangle &= -4B(B^2 - 1)(1 + n_{\mathrm{R}})(1 + n_{\mathrm{A}}) - 4F(1 - F^2)\, n_{\mathrm{R}} n_{\mathrm{A}}. \end{aligned} \tag{6.5}$$

In both lines, the first term comes from the product $-(B^2 - 1)$ of the BB entries of $(Q_0)_{\mathrm{RA}}$ and $(Q_0)_{\mathrm{AR}}$, and the second from the product $1 - F^2$ of their FF entries, with the sign of $\xi_3 \xi_2 \xi_1 \xi_4 = -n_{\mathrm{R}} n_{\mathrm{A}}$.

Let $\phi \in C_c^\infty(D^\circ)$, and set $\ell_1 = \phi\, Q_{\mathrm{RB,AB}}$ and $\ell_2 = \phi\, Q_{\mathrm{RB,AB}} Q^{\mathrm{FF}}_{\mathrm{RR}}$. These superfunctions have compact support in $\mathcal{O}^\circ$, and the invariance of the integral gives $J(\sigma \widehat{Y}_0 \ell_i) = \int \widehat{Y}_0 \ell_i = 0$. We compute $\langle \widehat{Y}_0 \ell_i \rangle_{\mathrm{top}}$ with the chain rule, (6.4), (6.5), and $Q^{\mathrm{FF}}_{\mathrm{RR}} = F + d\, n_{\mathrm{R}}$ and $Q^{\mathrm{BB}}_{\mathrm{AA}} - Q^{\mathrm{BB}}_{\mathrm{RR}} = -2B - d(n_{\mathrm{R}} + n_{\mathrm{A}})$ from Lemma 6.1(b). We obtain $\langle \widehat{Y}_0 \ell_i \rangle_{\mathrm{top}} = \alpha_i \partial_d \phi + \beta_i \partial_m \phi + \gamma_i \phi$, where

$$\begin{gathered} \alpha_1 = -(B^2 - F^2), \qquad \beta_1 = -2\big(B(B^2 - 1) + F(1 - F^2)\big), \qquad \gamma_1 = 0, \\ \alpha_2 = -F(B^2 - F^2) - d(B^2 - 1), \qquad \beta_2 = -2\big(B^2(B^2 - 1) + F^2(1 - F^2)\big), \qquad \gamma_2 = -d^2. \end{gathered}$$

For ℓ_2, the factor $F + d\, n_{\mathrm{R}}$ multiplies the top coefficient by F and the coefficient of n_{A} by d, and we used $F + d = B$.

We now return to the coordinates (B, F). Inserting the coordinate vector fields ∂_d and ∂_m computed above, we find $\alpha_i \partial_d \phi + \beta_i \partial_m \phi = V^{(i)} \cdot \nabla \phi$ with

$$V^{(1)} = \big(1 - B^2,\, 1 - F^2\big), \qquad V^{(2)} = \big(B(1 - B^2),\, F(1 - F^2)\big),$$

and therefore

$$0 = J\big(\sigma\, \widehat{Y}_0 \ell_i\big) = \int_D \frac{\sigma}{d^2} \big(V^{(i)} \cdot \nabla \phi + \gamma_i \phi\big)\, \mathrm{d}B\, \mathrm{d}F, \qquad i \in \{1, 2\}.$$

Finally, we conclude that σ is constant. A direct computation with $\partial_B d = 1$ and $\partial_F d = -1$ gives $\nabla \cdot (d^{-2} V^{(1)}) = 0$ and $\nabla \cdot (d^{-2} V^{(2)}) = -1 = d^{-2} \gamma_2$. Integrating by parts, we find $\int_D d^{-2} \phi\, V^{(i)} \cdot \nabla \sigma\, \mathrm{d}B\, \mathrm{d}F = 0$ for all ϕ, so that $V^{(i)} \cdot \nabla \sigma = 0$ on D° for $i = 1, 2$. The vectors $V^{(1)}$ and $V^{(2)}$ are linearly independent at every point of D°, since $\det(V^{(1)}, V^{(2)}) = (1 - B^2)(1 - F^2)(F - B) \neq 0$ there. Hence $\nabla \sigma = 0$, and $\sigma = c_1$ for a constant c_1 on the connected set $\mathcal{O}^\circ$, and by continuity on $\mathcal{O}$. This proves (6.3). □

At the corner, the odd part of the Jacobian degenerates, since $d = 0$ there, and the polar coordinates are singular at Λ. The following lemma is the corresponding integration formula. Its corner term is the boundary term of polar coordinates described in Section 3.6, and its bulk term is the classical form of the invariant measure in Efetov's coordinates.

Lemma 6.3. *Let f be a rapidly decreasing superfunction on $\mathcal{M}_1$. Then*

$$\int f = f_\varnothing(\Lambda) + \lim_{r \downarrow 0} \int_{\{p \in D : d > r\}} \frac{\langle f \rangle_{\mathrm{top}}(B, F)}{d^2}\, \mathrm{d}B\, \mathrm{d}F. \tag{6.6}$$

The bulk density $d^{-2} \langle f \rangle_{\mathrm{top}}$ is absolutely integrable on D, so the limit is the integral over D.

Proof. By Lemma 6.2, the identity (6.3) holds for all f with compact support in $\mathcal{O}$, with a constant c_1 that we determine at the end of the proof.

Now let f be general. Let $\psi \in C^\infty([0, \infty))$ be equal to one on $[0, 1]$ and to zero on $[2, \infty)$, and set $\psi_r = \psi(d_1 / r)$ for $r > 0$. In polar coordinates, we have $\psi_r = \psi(d/r)$, by Lemma 6.1(b), and ψ_r is invariant under $\mathfrak{k}$. We write $f = \psi_r f + (1 - \psi_r) f$.

The coefficients of the second term are supported where $d \geq r$. We justify extending (6.3) to this term by exhaustion with the cutoffs ψ_R of (5.21). On $\{B \geq 2\}$, choose body representatives $Q_0 = g \Lambda g^{-1}$ whose hyperbolic block and its inverse have entries bounded by $C B^{1/2}$, and whose spherical block is bounded in each of finitely many spherical charts. Then

$$Q = g\, e^{-g^{-1} \Theta_\xi g} \Lambda e^{g^{-1} \Theta_\xi g}\, g^{-1},$$

so the transition from polar coordinates to the translated regular chart centered at Q_0 has coefficients bounded by powers of B. The even coordinate shifts start at odd degree two, and there are four odd generators, so the polar top coefficient uses at most two body derivatives of the regular coefficients. Rapid decrease, the Cauchy–Schwarz inequality, and the polynomial volume bound of Lemma 5.1(c) therefore give an integrable majorant at infinity. The same majorant applies after multiplication by ψ_R, since its regular coefficients and their derivatives through order two are uniformly bounded in the translated charts for $R \geq 1$. On the remaining compact set away from the corner, we use Cartesian coordinates at the polar edges. Dominated convergence now applies on both sides of (6.3). Since $1 - \psi_r$ has no odd terms in polar coordinates, this gives

$$\int (1 - \psi_r) f = c_1 \int_D (1 - \psi(d/r)) \frac{\langle f \rangle_{\mathrm{top}}}{d^2} \, \mathrm{d}B \, \mathrm{d}F.$$

We now turn to the first term, which produces the corner term. Since ψ_r is invariant and compactly supported, Lemma 5.19 gives

$$\int \psi_r = \psi(0) = 1, \qquad \int \psi_r f = f_\emptyset(\Lambda) + \int \psi_r \big(f - f_\emptyset(\Lambda)\big).$$

We claim that the last integral tends to zero as $r \downarrow 0$. We use the exponential chart $Q = \Lambda e^{-2Y}$ of Step 3 of the proof of Lemma 5.19, in which $d_1 = \mathrm{STr}(Y^2) + O(Y^4)$, and we change variables by $Y = r^{1/2} Y'$, which has Berezinian one. The rescaled integrand is supported in a fixed compact set of Y_0', since the body of $\mathrm{STr}(Y_0'^2)$ is positive definite. Its coefficients converge to zero, since the body of $f - f_\emptyset(\Lambda)$ is evaluated at $r^{1/2} Y_0'$ and every term of positive degree in the odd coordinates comes with a positive power of $r^{1/2}$. Dominated convergence proves the claim.

To identify the smooth cutoff limit with the sharp cutoff in (6.6), note that Θ_ξ commutes with Λ, and hence

$$Q - \Lambda = e^{-\Theta_\xi} (Q_0 - \Lambda) e^{\Theta_\xi}.$$

Every coefficient on the right is $O(\sqrt{d})$ as $d \downarrow 0$, uniformly in the angles. Expressing the smooth function f in a fixed regular chart at Λ and using its finite Taylor expansion in the nilpotent coordinates gives $[f]_{\mathrm{top}} = O(\sqrt{d})$. In the coordinates $\delta = B - F$ and $s = B + F$, we have $\mathrm{d}B \, \mathrm{d}F = \frac{1}{2} \mathrm{d}\delta \, \mathrm{d}s$ and, for $0 < \delta < 1$, $2 - \delta \leq s \leq 2 + \delta$. Thus

$$\int_{\{d < \eta\}} d^{-2} \big| \langle f \rangle_{\mathrm{top}} \big| \, \mathrm{d}B \, \mathrm{d}F \leq C \int_0^\eta \delta^{-1/2} \, \mathrm{d}\delta \leq C \sqrt{\eta}, \qquad 0 < \eta < 1.$$

Together with the bound at infinity, this proves absolute integrability of the bulk density. Dominated convergence identifies both cutoff limits with its integral over D, and gives (6.6) with the constant c_1 in front of that integral.

It remains to show that $c_1 = 1$, and for this we use the invariance of the integral under an odd generator outside $\mathfrak{k}$. Let $Y = E_{\mathrm{RB,AF}}$ and $\phi \in C_c^\infty([0, \infty))$, and consider the

odd superfunction $\ell = \phi(d_1)\, Q_{\mathrm{RB,AF}}$, which has compact support. For an auxiliary odd generator ϑ, (5.6) and $[\Lambda, Y] = 2Y$ give

$$\vartheta \widehat{Y} d_1 = \tfrac{1}{2} \operatorname{STr}([\Lambda, \vartheta Y] Q) = \vartheta Q_{\mathrm{AF,RB}},$$
$$\vartheta \widehat{Y} Q_{\mathrm{RB,AF}} = [\vartheta Y, Q]_{\mathrm{RB,AF}} = \vartheta \big(Q_{\mathrm{AF,AF}} - Q_{\mathrm{RB,RB}} \big).$$

Since the integral is invariant, the Leibniz rule gives

$$0 = \int \widehat{Y} \ell = \int \phi'(d_1)\, Q_{\mathrm{AF,RB}} Q_{\mathrm{RB,AF}} + \int \phi(d_1) \big(Q_{\mathrm{AA}}^{\mathrm{FF}} - Q_{\mathrm{RR}}^{\mathrm{BB}} \big).$$

The superfunctions $\phi(d_1) Q_{\mathrm{RR}}^{\mathrm{BB}}$ and $\phi(d_1) Q_{\mathrm{AA}}^{\mathrm{FF}}$ are $\mathfrak{q}_{\mathrm{A}}$-closed, respectively $\mathfrak{q}_{\mathrm{R}}$-closed, by Lemma 5.18, and Lemma 5.19 evaluates the second integral as $\phi(0)(-1-1) = -2\phi(0)$. In the first integral, the integrand has vanishing body at Λ. Using Lemma 6.1(c) and the formula just proved, we find

$$\begin{aligned} \int \phi'(d_1) Q_{\mathrm{AF,RB}} Q_{\mathrm{RB,AF}} &= -c_1 \int_D \phi'(d) \frac{B^2 - F^2}{d^2} \, \mathrm{d}B \, \mathrm{d}F = -c_1 \int_0^\infty \phi'(\delta) \, \frac{1}{2\delta} \int_{|2-\delta|}^{2+\delta} s \, \mathrm{d}s \, \mathrm{d}\delta \\ &= -2c_1 \int_0^\infty \phi'(\delta) \, \mathrm{d}\delta = 2c_1 \phi(0), \end{aligned}$$

where we used the coordinates $\delta = B - F$ and $s = B + F$, in which $\mathrm{d}B \, \mathrm{d}F = \frac{1}{2} \mathrm{d}\delta \, \mathrm{d}s$ and $D = \{\delta \geq 0,\ |2 - \delta| \leq s \leq 2 + \delta\}$. Choosing $\phi(0) \neq 0$, we conclude that $c_1 = 1$. □

Remark 6.4. For a superfunction $f = \phi(B, F)$ without odd terms, the formula (6.6) reduces to $\int f = \phi(1,1)$, and the whole integral is the corner term. Without it, the integral would be zero, since the integrand has no top coefficient, and the formula $\int_D \phi \, d^{-2} \, \mathrm{d}B \, \mathrm{d}F$ that one might write instead diverges at the corner when $\phi(1,1) \neq 0$. This is the phenomenon of (3.16), and the identity $\int f = \phi(1,1)$ for invariant superfunctions is the form of the corner term in the physics literature [95, (2.69)]. The constant $c_1 = 1$ is also confirmed by the metallic limit. At $X = 0$, the formulas of Definition 1.4 give

$$K_X(x, y; z)\big|_{X=0} = \int_D e^{-zd} \, \mathrm{d}B \, \mathrm{d}F = \frac{1 - e^{-2z}}{z^2},$$

and I_X at $X = 0$ is the function I_{GUE} defined before Proposition 1.16.

6.2 The scalar semigroups

We now collect the properties of the semigroups on D that Definition 1.4 uses and that Part III needs. The radial diffusion $(\tilde{B}_t, \tilde{F}_t)$ is the image under $Q \mapsto (B, F)$ of the Brownian motion on $\mathcal{M}_1^0 = H_1 \times S_1$ with generator $X \Delta_{\mathcal{M}_1^0}$, since $\Delta_{\mathcal{M}_1^0}$ acts on functions of (B, F) as Δ_D and the diffusion is determined by its generator.

We now construct the singlet diffusion of Remark 1.5, which Part III uses throughout. We first recall the boundary behavior of the radial diffusion. The coordinate $\tilde{B}$ has the generator $X\partial_B((B^2-1)\partial_B)$, with scale density $(B^2-1)^{-1}$ and speed density X^{-1}. Its scale function

$$s(B) = \frac{1}{2}\log\frac{B-1}{B+1}$$

satisfies

$$s(1+) = -\infty, \qquad \int_1^2 |s(B)|\,\mathrm{d}B < \infty, \qquad \int_2^\infty \big(s(\infty) - s(B)\big)\,\mathrm{d}B = \infty.$$

By Feller's classification of boundary points [70], the first two properties make the point 1 an entrance boundary. The third shows that the point ∞ is not reached in finite time, and $\tilde{B}$ does not explode. The same computation, with scale density $(1-F^2)^{-1}$, shows that ± 1 are entrance boundaries for $\tilde{F}$. In particular, $(\tilde{B}_t, \tilde{F}_t) \in D^\circ$ for all $t \geq 0$ when $(\tilde{B}_0, \tilde{F}_0) \in D^\circ$. All starting points below lie in D° (the boundary of D has Lebesgue measure zero, and this restriction does not affect the kernels of Definition 1.4). We write $\mathbb{P}_p$ and $\mathbb{E}_p$ for the law of the radial diffusion started at p and its expectation.

The coordinates $\tilde{B}$ and $\tilde{F}$ have the generators

$$X(B^2-1)\partial_B^2 + 2XB\partial_B, \qquad X(1-F^2)\partial_F^2 - 2XF\partial_F,$$

and they stay in the open intervals $(1,\infty)$ and $(-1,1)$, on which these coefficients are smooth and the diffusion coefficients are positive. By a standard argument based on the martingale problem and Lévy's characterization [108], there exist independent standard Brownian motions β^1 and β^2, on the probability space of the radial diffusion, such that

$$\mathrm{d}\tilde{B}_t = 2X\tilde{B}_t\,\mathrm{d}t + \big(2X(\tilde{B}_t^2-1)\big)^{1/2}\,\mathrm{d}\beta_t^1, \qquad \mathrm{d}\tilde{F}_t = -2X\tilde{F}_t\,\mathrm{d}t + \big(2X(1-\tilde{F}_t^2)\big)^{1/2}\,\mathrm{d}\beta_t^2. \tag{6.7}$$

Since $\Delta_D d^{-1} = 0$ on D°, the function $1/d$ is positive and harmonic for the radial diffusion. For a positive function ϕ on D° with $\Delta_D\phi = 0$, we call the diffusion with generator $\phi^{-1}X\Delta_D(\phi\,\cdot)$, run until it leaves D°, the *Doob transform* of the radial diffusion by ϕ [105, 108]. For $\phi = 1/d$, the generator becomes $\phi^{-1}X\Delta_D(\phi\,\cdot) = Xd\Delta_D d^{-1} = X\Delta_D^{\mathrm{s}}$, by (6.1). Its drift is

$$2X\Big(B - \frac{B^2-1}{B-F}\Big)\partial_B + 2X\Big(-F + \frac{1-F^2}{B-F}\Big)\partial_F = 2X\,\frac{1-BF}{B-F}\,(\partial_B + \partial_F).$$

Since the two components of this drift coincide, the function d has no drift under the Doob transform. This is what makes the corner absorbing. We will use the elementary identity

$$(B-F)^2 - (1-BF)^2 = (B^2-1)(1-F^2) \geq 0, \qquad \text{so that} \qquad |1-BF| \leq B-F \quad \text{on } D. \tag{6.8}$$

For a stopping time τ and a process Z, we let Z^τ denote the process stopped at τ.

The general theory of Doob transforms constructs the transformed process up to the time at which it leaves D°, and it provides the formula (6.10) below for its law. What is specific to the domain D is that the transformed process leaves D° only through the corner, where we stop it, and that it reaches the corner with positive probability. This mass is not killed further, and it corresponds to the corner term of Lemma 6.3, as we explain after Lemma 6.7. The following lemma makes Remark 1.5 precise.

Lemma 6.5. *Let $X > 0$ and $p \in D^\circ$.*

(a) There exist a filtered probability space carrying a two-dimensional standard Brownian motion $(\beta^{s,1}, \beta^{s,2})$, a stopping time $\tau_c \in (0, \infty]$, and a continuous adapted process (B_t, F_t) with values in D, with the following properties. We have $(B_0, F_0) = p$, $(B_t, F_t) \in D^\circ$ for $t < \tau_c$, and $(B_t, F_t) = (1, 1)$ for $t \geq \tau_c$. On $[0, \tau_c)$,

$$\begin{aligned} \mathrm{d}B_t &= 2X\frac{1 - B_tF_t}{B_t - F_t}\,\mathrm{d}t + \bigl(2X(B_t^2 - 1)\bigr)^{1/2}\,\mathrm{d}\beta_t^{s,1}, \\ \mathrm{d}F_t &= 2X\frac{1 - B_tF_t}{B_t - F_t}\,\mathrm{d}t + \bigl(2X(1 - F_t^2)\bigr)^{1/2}\,\mathrm{d}\beta_t^{s,2}. \end{aligned} \tag{6.9}$$

Moreover, for all $T > 0$ and $n \geq 1$, there exists C such that $\mathbb{E}\sup_{t \leq T} B_t^n \leq CB_0^n$, where $p = (B_0, F_0)$.

(b) For all $t \geq 0$ and all nonnegative measurable functions Φ on the space of continuous paths $[0, t] \to D$,

$$\mathbb{E}\bigl[\Phi\bigl((B, F)|_{[0,t]}\bigr)\mathbf{1}_{t<\tau_c}\bigr] = \mathbb{E}_p\Bigl[\Phi\bigl((\tilde{B}, \tilde{F})|_{[0,t]}\bigr)\,\frac{d(p)}{d(\tilde{B}_t, \tilde{F}_t)}\Bigr]. \tag{6.10}$$

In particular, the law of $(B_t, F_t)_{t \geq 0}$ is determined by p and X. We write $\mathbb{E}_p^s$ for the expectation with respect to it.

(c) For all $z \in \mathbb{C}$ with $\operatorname{Re} z > 0$ and all $t \geq 0$,

$$\Psi_t^{X,z}(p) = \mathbb{E}_p^s \exp\Bigl(-z\int_0^t d(B_u, F_u)\,\mathrm{d}u\Bigr). \tag{6.11}$$

In particular, $|\Psi_t^{X,z}(p)| \leq \Psi_t^{X,\operatorname{Re} z}(p) \leq 1$, and $0 \leq \Psi_t^{X,z}(p) \leq 1$ for real $z > 0$.

The process (B_t, F_t) is the singlet diffusion of Remark 1.5, absorbed at the corner $(1, 1)$ at time τ_c, which we call its corner time. Started at the corner, it stays there. After the time τ_c, the integrand in (6.11) vanishes, and the mass that has reached the corner is not killed any further.

Proof. We begin with the Doob martingale. Let

$$M_t = \frac{d(p)}{d(\tilde{B}_t, \tilde{F}_t)}, \qquad t \geq 0.$$

The function $1/d$ is smooth and harmonic for Δ_D on D°, where the radial diffusion stays, and Itô's formula and (6.7) show that M is a positive continuous local martingale with $M_0 = 1$. For $m \geq 2$, let ϱ_m be the hitting time of $\{d \leq 1/m\}$ by the radial diffusion. Since the radial diffusion is continuous and never reaches the corner, $\varrho_m \to \infty$ almost surely, and the stopped process M^{ϱ_m} is a martingale bounded by $\max\{1, md(p)\}$. Indeed, if $d(p) \leq 1/m$, then $\varrho_m = 0$ and the stopped process equals one; otherwise continuity gives $d(\tilde{B}_{t\wedge\varrho_m}, \tilde{F}_{t\wedge\varrho_m}) \geq 1/m$.

The construction of Doob transforms [105, 108] uses these martingales as Girsanov densities, and we recall what it gives. There exist a filtered probability space carrying a two-dimensional standard Brownian motion $(\beta^{s,1}, \beta^{s,2})$ and a continuous adapted process (B_t, F_t) with $(B_0, F_0) = p$, defined up to the time $\zeta \in (0, \infty]$ at which it leaves D°, that solves (6.9) on $[0, \zeta)$, such that

$$\mathbb{E}\big[\Phi\big((B,F)|_{[0,\sigma]}\big)\mathbf{1}_{\sigma<\zeta}\big] = \mathbb{E}_p\big[\Phi\big((\tilde{B},\tilde{F})|_{[0,\sigma]}\big)M_\sigma\big] \tag{6.12}$$

for every bounded stopping time σ of the path, in particular for every fixed time, and all nonnegative measurable Φ. Indeed, Girsanov's theorem for the density $M_{\varrho_m\wedge T}$ adds to β^1 and β^2 the drifts $(2X(\tilde{B}^2-1))^{1/2}/d$ and $-(2X(1-\tilde{F}^2))^{1/2}/d$, up to the time $\varrho_m \wedge T$, and the identities

$$2B - \frac{2(B^2-1)}{d} = \frac{2(1-BF)}{d} = -2F + \frac{2(1-F^2)}{d}$$

turn (6.7) into (6.9). The coefficients of (6.9) are smooth on D°, so its solution is pathwise unique up to the exit time from every compact subset of D°, and unique in law. It remains to prove the properties that are specific to D.

We first show that the process leaves D° only through the corner. For $n \geq 2$, let

$$O_n = \big\{(B,F) : 1 + 1/n < B < n,\ |F| < 1 - 1/n\big\},$$

and let ϱ^s_m be the hitting time of $\{d \leq 1/m\}$ by (B, F). Fix m and T, and apply (6.12) with σ the minimum of T, of the exit time from O_n, and of the hitting time of $\{d \leq 1/m\}$, and with Φ the indicator of the event that the path leaves O_n before the two other times. On this event, $M_\sigma \leq md(p)$. Under $\mathbb{P}_p$, the probability that the radial diffusion leaves O_n before time T tends to zero as $n \to \infty$, because it stays in D° and does not explode. Therefore the probability that (B, F) leaves O_n before $\varrho^s_m \wedge T$ tends to zero as well. Since m, T, and n are arbitrary, (B, F) almost surely does not leave D° before ϱ^s_m, for every m. Let $\tau_c = \lim_{m\to\infty} \varrho^s_m$. If $\tau_c < \infty$, then $d(B_t, F_t)$ takes the values $1/m$ at times $\varrho^s_m \uparrow \tau_c$, and the path cannot stay in D° at the time τ_c. Therefore $\zeta = \tau_c$.

We next identify the limit at τ_c. The drifts of B and F in (6.9) coincide, and $d(B_t, F_t)$ is a nonnegative continuous local martingale on $[0, \tau_c)$. It converges as $t \uparrow \tau_c$ (a nonnegative local martingale is a supermartingale), and on the event $\{\tau_c < \infty\}$ its limit is zero, since

it takes the value $1/m$ at ϱ^{s}_m. As $d = (B-1) + (1-F)$ with both terms nonnegative, we conclude that $(B_t, F_t) \to (1,1)$. We set $(B_t, F_t) = (1,1)$ for $t \geq \tau_{\mathrm{c}}$. With $\zeta = \tau_{\mathrm{c}}$, the identity (6.12) is (6.10). Together with the absorption at the corner, it determines the law of (B, F), which proves (b).

For the moment bound, note that (6.8) bounds the drift in (6.9) by $2X$, and that the diffusion coefficient of B is at most $(2X)^{1/2}B$. We apply Itô's formula to B^n, stopped at $\varrho^{\mathrm{s}}_m \wedge T$. The Burkholder–Davis–Gundy inequality, Gronwall's inequality, and Fatou's lemma then yield

$$\mathbb{E} \sup_{t \leq T} B_t^n \leq C B_0^n,$$

where we used $B \geq 1$. This completes the proof of (a).

It remains to prove (c). Fix z with $\operatorname{Re} z > 0$, and set

$$\varphi(t) = \mathbb{E}^{\mathrm{s}}_p \exp\Big(- z \int_0^t d(B_u, F_u)\, \mathrm{d}u \Big).$$

The bound $d \leq 2B$ on D and the moment bound of part (a) allow us to differentiate under the expectation, and

$$\begin{aligned} \varphi'(t) &= -z\, \mathbb{E}^{\mathrm{s}}_p \Big[d(B_t, F_t) \exp\Big(- z \int_0^t d(B_u, F_u)\, \mathrm{d}u \Big) \Big] \\ &= -z\, d(p)\, \mathbb{E}_p \exp\Big(- z \int_0^t d(\tilde{B}_u, \tilde{F}_u)\, \mathrm{d}u \Big) = -z\, d(p) \big(e^{t(X\Delta_D - z d)} 1 \big)(p). \end{aligned}$$

In the second step, we used that $d(B_t, F_t) = 0$ for $t \geq \tau_{\mathrm{c}}$, and we applied (6.10), extended by linearity to integrable complex functionals, with

$$\Phi = d(B_t, F_t) \exp\Big(- z \int_0^t d(B_u, F_u)\, \mathrm{d}u \Big).$$

The integrability needed here follows from $d \leq 2B$ and the moment bound of part (a). The third step is the definition of the semigroup in Definition 1.4. Integrating from $\varphi(0) = 1$ and comparing with (1.5), we find $\varphi(t) = \Upsilon^{X,z}_t(p)$, which is (6.11). The bounds follow from $|\exp(-za)| = \exp(-a \operatorname{Re} z)$ for $a \geq 0$. □

For $X > 0$ and $z \in \mathbb{C}$, we let $\mathfrak{l}_z$ denote the sesquilinear form

$$\mathfrak{l}_z[h, h'] = X \int_D \Big((B^2 - 1)\, \partial_B h\, \overline{\partial_B h'} + (1 - F^2)\, \partial_F h\, \overline{\partial_F h'} \Big)\, \mathrm{d}B\, \mathrm{d}F + z \int_D d\, h\, \overline{h'}\, \mathrm{d}B\, \mathrm{d}F$$

on the space $C^\infty_{\mathrm{c}}(D)$ of smooth functions on D with compact support in D; the supports may meet the edges of D. By the theorem of Whitney on smooth even functions [136], the elements of $C^\infty_{\mathrm{c}}(D)$ are exactly the functions of (B, F) that are smooth and compactly supported as functions on $\mathcal{M}^0_1$. On D°, the singlet state is given by the Feynman–Kac formula (6.11) of Lemma 6.5(c), which implies the bounds in part (b) of the next lemma there, and part (b) extends them to D.

Lemma 6.6. *Let $X > 0$, let $z \in \mathbb{C}$ with $a = \operatorname{Re} z > 0$, and set $\Delta_{X,z} = X\Delta_D - zd$, so that $e^{t\Delta_{X,z}}$ is the semigroup of Definition 1.4.*

(a) The form $\mathfrak{l}_z$ is closable and sectorial on $L^2(D) = L^2(D, \mathrm{d}B\,\mathrm{d}F)$, and the semigroup of the associated operator agrees, on bounded functions in $L^2(D)$, with the Feynman–Kac semigroup $e^{t\Delta_{X,z}}$ of Definition 1.4. For bounded integrable functions f and g on D, we have $\int_D f\, e^{t\Delta_{X,z}} g = \int_D (e^{t\Delta_{X,z}} f)\, g$, and $|e^{t\Delta_{X,z}} f| \le e^{t\Delta_{X,a}} |f|$.

(b) For all $t \ge 0$, we have $|\Psi_t^{X,z}| \le \Psi_t^{X,a} \le 1$ on D and $\Psi_t^{X,z}(1,1) = 1$, and $\Psi_t^{X,a}$ is nonincreasing in t and in a.

(c) For all $0 < t_0 \le T$ and $m \ge 0$, there exists C such that, for $t \in [t_0, T]$ and $p \in D$,

$$\Psi_t^{X,a}(p) + \bigl(e^{t\Delta_{X,a}} 1\bigr)(p) \le C(1+B)^{-m}.$$

(d) The integrals defining $K_X(x,y;z)$, $\widetilde{K}_X(x;z)$, and $I_X(z)$ in Definition 1.4 converge absolutely for $0 < x \le y < 1$. The kernel K_X is bounded on $\{0 < x \le y < 1\}$, locally uniformly in z, and K_X, $\widetilde{K}_X$, and I_X are holomorphic in z in the half-plane $\operatorname{Re} z > 0$.

Proof. We begin with (a) for real $z = a$, which is standard. The operator $X\Delta_{\mathcal{M}_1^0}$ is essentially self-adjoint on the smooth compactly supported functions of the complete manifold $\mathcal{M}_1^0$, and its semigroup is the transition semigroup of the Brownian motion, since $\mathcal{M}_1^0$ has bounded geometry, which implies stochastic completeness [64]. Since the potential ad is nonnegative and continuous, the Trotter product formula gives

$$e^{t(X\Delta_{\mathcal{M}_1^0} - ad)} = \lim_{n\to\infty} \bigl(e^{tX\Delta_{\mathcal{M}_1^0}/n} e^{-tad/n}\bigr)^n$$

strongly, and the Feynman–Kac formula follows by dominated convergence in the corresponding Riemann sums [125].

Both semigroups commute with the rotations of H_1 and S_1, and on functions of (B, F) the space $L^2(\mathcal{M}_1^0)$ is $L^2(D)$, up to the factor $(2\pi)^2$, and the form of $-X\Delta_{\mathcal{M}_1^0} + ad$ is $\mathfrak{l}_a$ on $C_c^\infty(D)$. This proves the first assertion for real z, and the symmetry, the positivity, and the bound $e^{t\Delta_{X,a}} 1 \le 1$ follow. For complex z, the form $\mathfrak{l}_z$ is sectorial, has the same domain as $\mathfrak{l}_a$, and depends holomorphically on z. Both sides of the identification and of the symmetry are holomorphic in z, by [72] on the side of the forms and by dominated convergence on the side of the Feynman–Kac formula, and they agree for real z. The domination follows from $|\exp(-z\int d)| = \exp(-a\int d)$.

We next prove the bound (c) for the radial semigroup. Since $\tilde{F}_t \in [-1,1]$, we have $d(\tilde{B}_v, \tilde{F}_v) \ge \tilde{B}_v - 1$. By Itô's formula,

$$\log \tilde{B}_s = \log B + \int_0^s X(1 + \tilde{B}_v^{-2})\, \mathrm{d}v + \tilde{M}_s,$$

where $\tilde M$ is a continuous martingale with $\langle \tilde M\rangle_s \le 2Xs$. The drift is nonnegative, and $\tilde B_s \ge B e^{\tilde M_s}$. The exponential martingale inequality yields

$$\mathbb{P}\Big(\inf_{s\le T} \tilde M_s \le -\lambda \Big) \le e^{-\lambda^2/(4XT)}, \qquad \lambda \ge 0. \tag{6.13}$$

On the event that $\inf_{s\le T} \tilde M_s > -\frac12 \log B$, we have $\tilde B_s \ge \sqrt{B}$ for $s \le T$, and the killing up to time $t \ge t_0$ is at least $at_0(\sqrt{B}-1)$. Splitting the expectation along this event, we obtain

$$\big(e^{t\Delta_{X,a}}1\big)(p) \le e^{-at_0(\sqrt{B}-1)} + e^{-(\log B)^2/(16XT)},$$

which decays faster than every power of B.

We now prove (b). On D°, the bounds are those of Lemma 6.5(c), and the monotonicity follows from (6.11), since $d \ge 0$. To extend them to D, we show that $(t,p) \mapsto (e^{t\Delta_{X,z}}1)(p)$ is continuous on $[0,\infty)\times D$. Then $\Psi_t^{X,z}$ is continuous on D, since $|e^{v\Delta_{X,z}}1| \le 1$, and $\Psi_t^{X,z}(1,1) = 1$ since $d(1,1) = 0$.

Let first $z = a$ be real, and let $u_t = e^{-tA_{X,a}}1$, where $A_{X,a} = -X\Delta_{\mathcal{M}_1^0} + ad$ is realized as in Lemma 5.17(e); at rank one, $(d_1)_\varnothing = d$. By Lemma 5.17(e), applied with $f = 1$, the function $(t,Q) \mapsto u_t(Q)$ is smooth on $[0,\infty)\times \mathcal{M}_1^0$. The cutoffs ψ_R of (5.21) are compactly supported functions of B, and by (a) and the consistency of the realizations, $e^{-tA_{X,a}}\psi_R$ is the Feynman–Kac function $e^{t\Delta_{X,a}}\psi_R$. As $R\to\infty$, the first converges to u_t in L^2_{-q} for $q > \frac12$, and the second converges pointwise to $e^{t\Delta_{X,a}}1$, by dominated convergence. Hence the two functions agree almost everywhere. Since u_t is continuous, it follows that $|u_t| \le 1$ and that u_t is invariant under the rotations of H_1 and S_1, and we regard it as a function on D. For $0 < \delta < t$, the Markov property of the radial diffusion gives

$$\big(e^{t\Delta_{X,a}}1\big)(p) = \mathbb{E}_p\Big[e^{-a\int_0^\delta d(\tilde B_v,\tilde F_v)\,\mathrm{d}v}\, u_{t-\delta}(\tilde B_\delta, \tilde F_\delta)\Big],$$

where we replaced $e^{(t-\delta)\Delta_{X,a}}1$ by $u_{t-\delta}$, which is possible since the law of $(\tilde B_\delta, \tilde F_\delta)$ is absolutely continuous, the Brownian motion on $\mathcal{M}_1^0$ having a heat kernel [64]. As $\delta \downarrow 0$, the right side tends to $u_t(p)$, by the continuity of u and dominated convergence. Hence $e^{t\Delta_{X,a}}1 = u_t$ everywhere, and this function is smooth on $[0,\infty)\times D$, by the theorem of Whitney used in Lemma 6.7(a).

For complex z, let $(t_n,p_n) \to (t,p)$ in $[0,\infty)\times D$. The functions $z \mapsto (e^{t_n\Delta_{X,z}}1)(p_n)$ are holomorphic in the half-plane $\operatorname{Re} z > 0$, by dominated convergence in the Feynman–Kac formula, and bounded by one, and they converge for real $z > 0$, by the previous paragraph. By Vitali's theorem, they converge for every z in the half-plane, and the limit is the holomorphic function $z \mapsto (e^{t\Delta_{X,z}}1)(p)$, since the two agree for real z.

It remains to prove the bound (c) for $\Psi_t^{X,a}$, and we argue as for the radial diffusion. By Itô's formula, we have, before τ_c,

$$\log B_s = \log B + \int_0^s \gamma(B_v, F_v)\,\mathrm{d}v + M'_s,$$
$$\gamma = X\Big(1 + \frac{1}{B^2} - \frac{2(B^2-1)}{dB}\Big) \in [-3X, 2X],$$

where $\langle M'\rangle_s \le 2Xs$, and we used $d \ge B-1$, so that $(B^2-1)/(dB) \le (B+1)/B \le 2$. On the event that $\inf_{s\le T} M'_s > -\frac{1}{2}\log B$, we have $B_s \ge \sqrt{B}\,e^{-3XT}$ for $s \le T$. For $B \ge 4e^{6XT}$, the process then does not reach the corner, and the killing up to time $t \ge t_0$ is at least $at_0(\sqrt{B}e^{-3XT}-1)$. This yields

$$\Psi_t^{X,a}(p) \le e^{-at_0(\sqrt{B}e^{-3XT}-1)} + e^{-(\log B)^2/(16XT)}, \qquad B \ge 4e^{6XT},$$

and $\Psi_t^{X,a} \le 1$ elsewhere.

Finally, we prove (d). Fix $0 < x \le y < 1$ and set $\ell = y - x$. Using (a) and (b), we bound the integrand of K_X by $\Psi_x^{X,a}\, e^{\ell\Delta_{X,a}}\Psi_{1-y}^{X,a}$.

One of the three times x, ℓ, and $1-y$ is at least $\frac{1}{3}$. If $x \ge \frac{1}{3}$, we bound $e^{\ell\Delta_{X,a}}\Psi_{1-y}^{X,a} \le 1$ and use (c) for $\Psi_x^{X,a}$. For $1-y \ge \frac{1}{3}$, we first move the semigroup to the left factor by the symmetry in (a). For $\ell \ge \frac{1}{3}$, we bound the two singlet states by one and use (c) for $e^{\ell\Delta_{X,a}}1$. In all cases, the integral is bounded by

$$C\int_D (1+B)^{-3}\,\mathrm{d}B\,\mathrm{d}F < \infty,$$

with C uniform over the triangle and over $\mathrm{Re}\,z$ in compact subsets of $(0,\infty)$. The factor B/d in $\widetilde{K}_X$ is integrable near the corner and bounded by two for $B \ge 3$, so the absolute convergence of $\widetilde{K}_X$ follows from (c).

We now turn to holomorphy. Dominated convergence in the Feynman–Kac formula shows that $z \mapsto (e^{t\Delta_{X,z}}f)(p)$ is holomorphic for bounded f and each p, and so are $\Psi_t^{X,z}(p)$ and $(e^{\ell\Delta_{X,z}}\Psi_{1-y}^{X,z})(p)$. The domination above is locally uniform in z, since $\Psi_t^{X,a}$ and $e^{t\Delta_{X,a}}1$ are nonincreasing in a. By Morera's theorem and Fubini's theorem, K_X and $\widetilde{K}_X$ are holomorphic, and the holomorphy of I_X follows from the uniform bound on the triangle. □

The argument for Lemma 6.6(c), in the scaled variables of Section 10.2, gives the tail envelope of Lemma 10.3, which is uniform in the damping and which Section 10 uses for the residues at zero damping.

6.3 Invariant superfunctions and the singlet state

We now show that the semigroup $e^{-t\mathsf{G}_{X,z}}$, acting on superfunctions invariant under $\mathfrak{k}$, is the singlet semigroup on D, and that the quotient $(\phi - \phi(1,1))/d$ of a radial representative

ϕ evolves by the radial semigroup. The main input is the radial part of C_1, which we compute in Lemma 6.8 after the proof of the first two parts of the following lemma.

Lemma 6.7. *(a) Let f be a smooth superfunction on $\mathcal{M}_1$ that is invariant under $\mathfrak{k}$. Then, in polar coordinates, $f = \phi(B, F)$ has no odd terms and does not depend on the angles, and ϕ is a smooth function on D with $\phi(1,1) = f_\varnothing(\Lambda)$. We call ϕ the* radial representative *of f.*

(b) Let $X > 0$ and $z > 0$. Then

$$\big(e^{-t\mathsf{G}_{X,z}}1\big)_\varnothing(\Lambda) = 1, \qquad t \geq 0. \tag{6.14}$$

(c) Let $X > 0$ and $z > 0$, and let f be a rapidly decreasing superfunction invariant under $\mathfrak{k}$, with radial representative ϕ, and let $c = \phi(1,1)$. Then $(\phi - c)/d$ is bounded and belongs to $L^2(D)$, and for $t \geq 0$ the radial representative of $e^{-t\mathsf{G}_{X,z}} f$ is

$$c\,\Psi_t^{X,z} + d\,e^{t(X\Delta_D - zd)}\Big(\frac{\phi - c}{d}\Big) = \mathbb{E}_p^{\mathrm{s}}\Big[\phi(B_t, F_t)\exp\Big(-z\int_0^t d(B_v, F_v)\,\mathrm{d}v\Big)\Big]. \tag{6.15}$$

The expectation is taken for $p \in D^\circ$; the algebraic expression gives its continuous extension to D for $t > 0$, with value c at the corner. Thus $e^{-t\mathsf{G}_{X,z}}$ acts on invariant superfunctions as the semigroup of the singlet diffusion, absorbed at the corner and killed at rate zd.

(d) Let $X > 0$ and $z > 0$. For $t \geq 0$, the radial representative of $e^{-t\mathsf{G}_{X,z}}1$ is the singlet state $\Psi_t^{X,z}$ of Definition 1.4.

Proof of Lemma 6.7(a) and (b). We begin with (a). Write $Q = e^{-\Theta_\xi} Q_0 e^{\Theta_\xi}$ as in (6.2), and let

$$f(s) = f\big(e^{-s\Theta_\xi} Q_0 e^{s\Theta_\xi}\big), \qquad s \in [0,1],$$

an expression in the Grassmann algebra generated by the ξ_a. Since $\Theta_\xi = \sum_a \xi_a Y_a$ with $Y_a \in \mathfrak{k}_1$, (5.6) and the chain rule give

$$\frac{\mathrm{d}}{\mathrm{d}s} f(s) = -\sum_a \xi_a (\widehat{Y}_a f)(e^{-s\Theta_\xi} Q_0 e^{s\Theta_\xi}) = 0.$$

It follows that $f = f(0) = f_\varnothing(Q_0)$ in polar coordinates. Since the rotations e^{sY} with $Y \in \mathfrak{k}_0$ act transitively on the angles θ_{B} and θ_{F}, the body $f_\varnothing(Q_0)$ does not depend on them. It is a smooth function on $\mathcal{M}_1^0$ that is invariant under the rotations of H_1 and S_1, hence a smooth function of $(B, F) \in D$ by [136], since $B - 1$ and $1 - F$ are smooth functions of the squared distances to the base points with nonvanishing derivatives. Finally, $Q_0 = \Lambda$ at $(B, F) = (1, 1)$.

We next prove (b). For $t > 0$, the superfunction $e^{-t\mathsf{G}_{X,z}}1$ is rapidly decreasing, by Proposition 5.15(b), and invariant under $\mathfrak{k}$, by Corollary 5.20. By Lemma 5.18, we have $\mathfrak{q}_{\mathrm{A}} e^{-t\mathsf{G}_{X,z}}1 = 0$, and Lemma 5.19 and Corollary 5.20 give

$$\big(e^{-t\mathsf{G}_{X,z}}1\big)_\varnothing(\Lambda) = \int e^{-t\mathsf{G}_{X,z}}1 = 1.$$

The case $t = 0$ is trivial. □

The other two parts of Lemma 6.7 rest on the following lemma, which computes the radial part of C_1, that is, the operator that C_1 induces on radial representatives. We set $A = \operatorname{diag}(B^2 - 1, 1 - F^2)$, so that $\Delta_D = \nabla \cdot (A\nabla)$.

Lemma 6.8. *Let f be a smooth superfunction on $\mathcal{M}_1$ that is invariant under $\mathfrak{k}$, with radial representative ϕ. Then the superfunction $\mathsf{C}_1 f$ is invariant under $\mathfrak{k}$, and its radial representative is $-\Delta_D^{\mathrm{s}}\phi$ on D minus the corner. Moreover,*

$$\Delta_D^{\mathrm{s}} = d^2\,\nabla \cdot \big(d^{-2} A \nabla\big) = d^2\Big(\partial_B\big(d^{-2}(B^2-1)\partial_B\big) + \partial_F\big(d^{-2}(1-F^2)\partial_F\big)\Big). \tag{6.16}$$

Proof. The superfunction $\mathsf{C}_1 f$ is invariant by Lemma 5.8(a). Using Lemma 6.7(a), we see that C_1 induces on radial representatives a linear operator L, which is a differential operator of order at most two in (B, F) on D°, by Lemma 5.8(c), and $L1 = 0$. We determine L with Lemma 5.11: its second-order part comes from the body, and its first-order part from the values of C_1 on two invariant polynomials in the entries of Q.

We begin with the second-order part. If u and v are even and invariant, then so is $\mathcal{Q}_{\mathsf{C}_1}(u, v)$, by its definition and Lemma 5.8(a), and its radial representative is its body at the body points Q_0, by Lemma 6.7(a). By Lemma 5.11(a), this body is $-2\nabla u_\emptyset \cdot \nabla v_\emptyset$. On functions of (B, F), the metric (5.2) has $|\nabla B|^2 = B^2 - 1$, $|\nabla F|^2 = 1 - F^2$, and $\nabla B \cdot \nabla F = 0$, by Section 5.2. If u and v have the radial representatives ϕ_u and ϕ_v, we conclude that $\mathcal{Q}_{\mathsf{C}_1}(u, v)$ has the radial representative $-2A\nabla\phi_u \cdot \nabla\phi_v$.

Next, we write f as a function of two invariant superfunctions. By Lemma 6.1(b), the invariant superfunctions d_1 and $\frac{1}{4}\operatorname{STr}((\Lambda Q)^2)$ have no odd terms in polar coordinates, and their radial representatives are d and $m = B^2 - F^2$. These are coordinates on D°, since $\partial(d, m)/\partial(B, F) = 2d \neq 0$, and we regard ϕ as a smooth function of (d, m) on D°. Then $f = \phi(d_1, \frac{1}{4}\operatorname{STr}((\Lambda Q)^2))$ on the set $\{B > 1, |F| < 1\}$, since both sides are invariant with the radial representative ϕ. For a smooth function ϕ of two variables and even superfunctions u_1 and u_2, the chain rule of Lemma 3.4 and the definition of $\mathcal{Q}_{\mathsf{C}_1}$ give

$$\mathsf{C}_1\phi(u_1, u_2) = \sum_a \partial_a\phi(u_1, u_2)\,\mathsf{C}_1 u_a + \frac{1}{2}\sum_{a,b}\partial_a\partial_b\phi(u_1, u_2)\,\mathcal{Q}_{\mathsf{C}_1}(u_a, u_b).$$

We now reduce L to two values. We apply the last identity with $u_1 = d_1$ and $u_2 = \frac{1}{4}\operatorname{STr}((\Lambda Q)^2)$. By the computation of the second-order part, the second sum has the same radial representative as the second-order part of $-\Delta_D\phi$. Since $-\Delta_D^{\mathrm{s}}$ differs from $-\Delta_D$ by a first-order operator, by (6.1), it follows that $L = -\Delta_D^{\mathrm{s}}$ on D° if $Ld = -\Delta_D^{\mathrm{s}} d$ and $Lm = -\Delta_D^{\mathrm{s}} m$. Using $\Delta_D^{\mathrm{s}}(dh) = d\,\Delta_D h$ with $h = 1$ and $h = B + F$, we find $\Delta_D^{\mathrm{s}} d = 0$ and $\Delta_D^{\mathrm{s}} m = 2d^2$. By Lemma 5.11(b) with $k = 1$, we have $Ld = 0$ and $Lm = -2d^2$, and therefore $L = -\Delta_D^{\mathrm{s}}$ on D°. Both sides are continuous up to the edges of D minus the corner, which proves the first assertion.

Finally, (6.16) follows from the second form of Δ_D^{s} in (6.1), as one checks with $\partial_B d^{-2} = -2d^{-3}$ and $\partial_F d^{-2} = 2d^{-3}$. □

The identity (6.16) is the general formula for the radial part of an invariant Laplace operator, with the radial density d^{-2} of Lemma 6.3 [39], and it shows that Δ_D^{s} is symmetric for the measure $d^{-2}\,\mathrm{d}B\,\mathrm{d}F$. We now prove the remaining parts of Lemma 6.7.

Proof of Lemma 6.7(c) and (d). We first identify the stochastic expression in (6.15). Let f, ϕ, and c be as in (c), and let $f_t = e^{-t\mathsf{G}_{X,z}} f$. The superfunction f_t is invariant under $\mathfrak{k}$, by the argument of the last assertion of Corollary 5.20, which applies to rapidly decreasing superfunctions. By Lemma 5.17(d), the coefficients of f_t are smooth functions of $t \in [0,\infty)$ and of the body point, and for all i, the norms of $\nabla^i f_t$ in all the weighted spaces of Section 5.2 are bounded uniformly in $t \in [0,T]$. Let ϕ_t denote the radial representative of f_t. By (a), $(t,p) \mapsto \phi_t(p)$ is smooth on $[0,\infty) \times D$, and its derivatives are bounded on compact subsets of D, uniformly in $t \in [0,T]$.

We first compute the corner value. Applying Lemma 5.19 to the $\mathfrak{q}_{\mathrm{A}}$-closed superfunction f_t, and Proposition 5.15(c) with $1 \in L^2_{-q}$, we find

$$\phi_t(1,1) = \int f_t = \mathcal{B}_1(1, e^{-t\mathsf{G}_{X,z}} f) = \mathcal{B}_1(e^{-t\mathsf{G}_{X,z}} 1, f) = \int (e^{-t\mathsf{G}_{X,z}} 1) f.$$

By Lemma 5.18, we have $\mathfrak{q}_{\mathrm{A}}\big((e^{-t\mathsf{G}_{X,z}} 1) f\big) = 0$. This superfunction has polynomially integrable coefficients, and Lemma 5.19 and (6.14) evaluate the right side as $\phi(1,1)$. Thus $\phi_t(1,1) = c$ for all $t \geq 0$.

We use the singlet diffusion of Lemma 6.5 to identify ϕ_t directly. First, the uniform rapid bounds above and Sobolev embedding through order three on the four-dimensional body $H_1 \times S_1$, in its translated charts of bounded geometry, give

$$C_T := \sup_{0 \leq r \leq T} \sup_{p \in D} |\phi_r(p)| < \infty.$$

By Lemma 6.8,

$$\partial_r \phi_r = X \Delta_D^{\mathrm{s}} \phi_r - z d \phi_r \qquad \text{on } D^\circ. \tag{6.17}$$

Fix $T > 0$ and $p \in D^\circ$, and write $Z_s = (B_s, F_s)$ for the singlet diffusion started at p. For all sufficiently large n let σ_n be its exit time from

$$O_n = \{(B,F) : 1 + 1/n < B < n,\ |F| < 1 - 1/n\}.$$

These are increasing relatively compact open subsets exhausting D°. By Lemma 6.5(a), the paths do not explode or reach another edge before the corner time, and $\sigma_n \uparrow \tau_{\mathrm{c}}$ almost surely.

Apply Itô's formula before $T \wedge \sigma_n$ to

$$M_s = \exp\Big(-z \int_0^s d(Z_v)\,\mathrm{d}v\Big) \phi_{T-s}(Z_s).$$

Equation (6.17) cancels its drift. On $\overline{O}_n$ the relevant coefficients and derivatives are bounded, so the stopped stochastic integral is a true martingale. Moreover $|M_{s\wedge\sigma_n}| \le C_T$ for $0 \le s \le T$, because $z > 0$. Thus

$$\phi_T(p) = \mathbb{E}_p^{\mathrm{s}}\Big[e^{-z\int_0^{T\wedge\sigma_n} d(Z_v)\,\mathrm{d}v}\phi_{T-(T\wedge\sigma_n)}(Z_{T\wedge\sigma_n})\Big].$$

Let $n \to \infty$. On $\{\tau_{\mathrm{c}} \le T\}$, joint continuity and the preserved corner value give the limit c for the last factor, including when $\tau_{\mathrm{c}} = T$. After absorption $Z_s = (1,1)$ and $d(Z_s) = 0$, so this is exactly the terminal value $\phi(Z_T)$ with the discount continued to time T. On $\{T < \tau_{\mathrm{c}}\}$ the limit is the usual terminal value at T. Bounded convergence therefore gives

$$\phi_T(p) = \mathbb{E}_p^{\mathrm{s}}\Big[\phi(Z_T)\exp\Big(-z\int_0^T d(Z_v)\,\mathrm{d}v\Big)\Big]. \tag{6.18}$$

We now identify the algebraic expression in (6.15). Let φ be a bounded measurable function on D and $p \in D^\circ$. Paths absorbed by time t contribute zero to the following expectation, since d vanishes at the corner. Apply (6.10) to the functional

$$\Phi(\omega) = (d\varphi)(\omega_t)\exp\Big(-z\int_0^t d(\omega_v)\,\mathrm{d}v\Big)$$

of a continuous path $\omega\colon [0,t] \to D$, first for nonnegative φ and then by linearity for complex φ. The moment bound of Lemma 6.5(a), $d \le 2B$, and the boundedness of φ justify integrability on the singlet side; on the radial side the factor d cancels and the integrand is bounded. We obtain

$$\mathbb{E}_p^{\mathrm{s}}\Big[(d\varphi)(B_t,F_t)\,e^{-z\int_0^t d(B_v,F_v)\,\mathrm{d}v}\Big] = d(p)\big(e^{t(X\Delta_D - zd)}\varphi\big)(p). \tag{6.19}$$

Set $g = (\phi - c)/d$ away from the corner and assign it any finite value there. Smoothness of ϕ and $\phi(1,1) = c$ give $|\phi - c| \le Cd$ near the corner. For $B \ge 3$, boundedness of ϕ and $d \ge B - 1$ give $|g| \le C/B$. Hence $g \in L^\infty(D) \cap L^2(D)$. In this application $dg = \phi - c$ is itself bounded. Splitting $\phi = c + dg$ in (6.18), and applying (6.19) and (6.11), yields

$$\phi_t(p) = c\Psi_t^{X,z}(p) + d(p)\big(e^{t(X\Delta_D - zd)}g\big)(p), \qquad p \in D^\circ.$$

For $t > 0$, the scalar smoothing assertion Lemma 5.17(e), applied to $g \in L^2(D)$, makes the radial semigroup term smooth on the underlying manifold, hence continuous up to the edges in radial coordinates. At the corner, its product with d tends to zero by $|e^{t(X\Delta_D - zd)}g| \le \|g\|_\infty$, while $\Psi_t^{X,z}$ is continuous and equals one there by Lemma 6.6(b). Thus the algebraic identity extends to all of D; at $t = 0$ it is immediate. This proves (c).

We now prove (d). Let $0 < t_0 < t$. We apply what we have proved to the rapidly decreasing invariant superfunction $e^{-t_0\mathsf{G}_{X,z}}1$, whose radial representative we denote by

ϕ_{t_0} and whose corner value is 1, by (6.14). Since $e^{-t\mathsf{G}_{X,z}}1 = e^{-(t-t_0)\mathsf{G}_{X,z}}e^{-t_0\mathsf{G}_{X,z}}1$, the radial representative of $e^{-t\mathsf{G}_{X,z}}1$ is

$$\Psi^{X,z}_{t-t_0} + d\, e^{(t-t_0)\Delta_{X,z}} g_{t_0}, \qquad g_{t_0} = \frac{\phi_{t_0} - 1}{d}.$$

We let $t_0 \downarrow 0$. By Lemma 5.17(c) with $f = 1$, the body of $e^{-t_0\mathsf{G}_{X,z}}1$ and its derivatives of order at most two converge to those of the constant 1 as $t_0 \downarrow 0$, uniformly on compact subsets of $\mathcal{M}_1^0$. Near Λ, this body minus one is invariant under the rotations about Λ, so that its differential vanishes at Λ, and Taylor's formula bounds it by its second derivatives times the squared distance to Λ, which is at most Cd. Hence $g_{t_0} \to 0$ uniformly on compact subsets of D, and $|g_{t_0}| \le C$ near the corner, uniformly in $t_0 \le t$. Away from the corner, the weighted bound of Lemma 5.17(c) and local Sobolev bounds imply that $|g_{t_0}| \le C(1+B)^K$ for some C and K and all $t_0 \le t$.

In the Feynman–Kac formula

$$\big(e^{(t-t_0)\Delta_{X,z}} g_{t_0}\big)(p) = \mathbb{E}_p\Big[g_{t_0}(\tilde B_{t-t_0}, \tilde F_{t-t_0})\, e^{-z\int_0^{t-t_0} d(\tilde B_v, \tilde F_v)\,\mathrm{d}v}\Big],$$

the integrand therefore tends to zero almost surely, by the continuity of the paths, and it is bounded by $C(1+\sup_{v\le t}\tilde B_v)^K$. This bound is integrable, since $\tilde B_v \le Be^{2Xv+\tilde M_v}$ with $\tilde M$ as in the proof of Lemma 6.6, and by the exponential martingale inequality applied to $-\tilde M$. Dominated convergence and the continuity of $t \mapsto \Psi^{X,z}_t$ show that the radial representative of $e^{-t\mathsf{G}_{X,z}}1$ is $\Psi^{X,z}_t$, which proves (d). □

Remark 6.9. Two supplementary scripts confirm, by exact computer algebra in the Grassmann algebra with four generators, the computations written out in Sections 6.1 and 6.3. They use only (5.6), (5.14), and (6.2). In the proof of Lemma 6.2, we showed that σ is constant, so that J is invariant under all of $\mathfrak{g}$. In polar coordinates, with $\vartheta_Y\widehat{Y} = \sum_x v^x\partial_x + \sum_a v^{\xi_a}\partial_{\xi_a}$, where x runs over B, F, θ_{B}, and θ_{F}, this invariance is the divergence identity

$$\sum_a \partial_{\xi_a}\big(d^{-2}v^{\xi_a}\big) - \sum_x \partial_x\big(d^{-2}v^x\big) = 0, \tag{6.20}$$

with left derivatives in the ξ_a. This is the classical form $d^{-2}\,\mathrm{d}B\,\mathrm{d}F\,\mathrm{d}\theta_{\mathrm{B}}\,\mathrm{d}\theta_{\mathrm{F}}\,\mathrm{d}\xi$ of the invariant measure in Efetov's coordinates, which [95, (2.67)] states for a parametrization that differs from ours by a linear change of the odd variables. The script `polar/check_polar_density.py` confirms (6.20) for the sixteen generators E_{ij}, and shows that it fails if d^{-2} is replaced by d^{-1} or d^{-3}. The script `polar/test_singlet.py` confirms Lemma 6.8. It normalizes the auxiliary operator of `casimir.py` by $-\frac{1}{2}$ to obtain C_1, applies it to the invariant superfunctions with radial representatives B, F, B^2, BF, F^2, and $e^B\cos F$ at two rational points of D°, and asserts exact equality with $-\Delta_D^{\mathrm{s}}$ in every Grassmann coefficient. The radial part (6.16) is symmetric for $d^{-2}\,\mathrm{d}B\,\mathrm{d}F$, in agreement with the density.

Lemma 10.7 computes the operator C_1 in the polar chart on all superfunctions on $\{B > 1\}$, not only on invariant ones; Section 10 needs it for the semigroup products with sources of Definition 10.6, which are not invariant.

Lemma 6.7 explains the two terms of the singlet state. For $d > 0$, define $h_t = (1 - \Psi_t^{X,z})/d$; by (1.5), $h_t = z \int_0^t e^{s(X\Delta_D - zd)} 1 \, \mathrm{d}s$. The radial representative of the supersymmetric semigroup applied to 1 is $\Psi_t^{X,z}$, by Lemma 6.7(d), and its corner value is preserved by (6.14). In probabilistic terms, the singlet diffusion is the radial diffusion conditioned by the harmonic function $1/d$, which blows up at the corner, so it is attracted by the corner and absorbed there, and the absorbed mass is not killed further since $d = 0$ at the corner. The mass that it carries to the corner is the probabilistic counterpart of the corner term in Lemma 6.3. It is the reason why $\Psi_t^{X,z}(1,1) = 1$ although the radial semigroup decays at every point. Figure 6.1 illustrates the two diffusions.

The function $1/d$ also has a meaning on $\mathcal{M}_1^0$, which the following lemma records; we do not use it in the sequel. The lemma says that $1/d$ is $8\pi^2$ times the Green function of $-\Delta_{\mathcal{M}_1^0}$ with pole at Λ, so that the singlet diffusion is the Brownian motion of $\mathcal{M}_1^0$, seen through the coordinates (B, F), conditioned to converge to the point Λ in the sense of Doob [105].

Lemma 6.10. *The function $1/d$ is locally integrable on $\mathcal{M}_1^0$, and*

$$\int_{\mathcal{M}_1^0} \frac{1}{d} \, \Delta_{\mathcal{M}_1^0} \phi \, \mathrm{dvol} = -8\pi^2 \phi(\Lambda) \tag{6.21}$$

for all smooth compactly supported functions ϕ on $\mathcal{M}_1^0$. Moreover, $1/d = 8\pi^2 G$, where G is the minimal positive Green function of $-\Delta_{\mathcal{M}_1^0}$ with pole at Λ.

Proof. The rotations of H_1 and S_1 about their base points fix Λ, preserve d and dvol, and commute with $\Delta_{\mathcal{M}_1^0}$. On functions of (B, F), the measure dvol is $(2\pi)^2 \, \mathrm{d}B \, \mathrm{d}F$, as in the proof of Lemma 6.6. In the coordinates $\delta = B - F$ and $s = B + F$ of the proof of Lemma 6.3, we have $\int_{\{d<1\}} d^{-1} \, \mathrm{d}B \, \mathrm{d}F = \int_0^1 \delta^{-1} \delta \, \mathrm{d}\delta = 1$, and $1/d$ is bounded on $\{d \geq 1\}$, so $1/d$ is locally integrable.

We now prove (6.21). By the invariance under the rotations, we may replace ϕ by its average over them, which is a function of (B, F). The left side of (6.21) is then $(2\pi)^2$ times the limit as $r \downarrow 0$ of $\int_{\{d>r\}} d^{-1} \Delta_D \phi \, \mathrm{d}B \, \mathrm{d}F$. Recall the matrix A, defined before Lemma 6.8, with $\Delta_D = \nabla \cdot (A\nabla)$, and let ν be the unit normal of $\{d > r\}$ that points toward the corner. Since $\Delta_D d^{-1} = 0$, Green's formula gives

$$\int_{\{d>r\}} \frac{\Delta_D \phi}{d} \, \mathrm{d}B \, \mathrm{d}F = \int_{\{d=r\}} \Big(\frac{1}{d} \, A\nabla\phi - \phi \, A\nabla \frac{1}{d} \Big) \cdot \nu \, \mathrm{d}s, \qquad A\nabla \frac{1}{d} \cdot \nu = \frac{B+F}{\sqrt{2}\, d},$$

where the edges of D contribute nothing, since $A\nu$ vanishes there. The segment $\{d = r\}$ has length $\sqrt{2}\, r$, and $A\nabla\phi = O(r)$ on it, so the right side tends to $-2\phi(1,1)$. This proves (6.21), which says that $-\Delta_{\mathcal{M}_1^0}(1/d) = 8\pi^2 \delta_\Lambda$ in the sense of distributions.

Next, let G_n be the Dirichlet Green function of $-\Delta_{\mathcal{M}_1^0}$ with pole at Λ in $\Omega_n = \{B < n\}$. The minimal positive Green function G is the increasing limit of the G_n [64]. By (6.21) and Weyl's lemma, the function $1/d - 8\pi^2 G_n$ is harmonic in Ω_n. Its boundary values $1/d$ on $\{B = n\}$ lie in $[(n+1)^{-1}, (n-1)^{-1}]$, and by the maximum principle, we have $0 \le 1/d - 8\pi^2 G_n \le (n-1)^{-1}$ on Ω_n. Letting $n \to \infty$, we obtain $1/d = 8\pi^2 G$. □

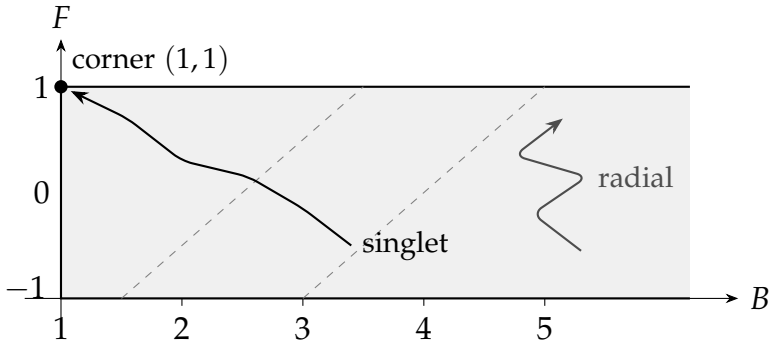


Figure 6.1. The domain $D = [1,\infty) \times [-1,1]$. The dashed lines are level lines of $d = B - F$, which vanishes only at the corner. The radial diffusion, with generator $X\Delta_D$, does not reach the boundary of D at positive times (the edges are entrance boundaries), and it is killed at rate zd. The singlet diffusion, its Doob transform by $1/d$, is drawn toward the corner and is absorbed there without further killing. The singlet state $\Psi_t^{X,z}$ is the expectation of the survival factor of the singlet diffusion.

6.4 The rank-one calculus

We now develop the calculus of pairings at rank one that Part III uses. It evaluates the mixed pairings of singlet states with covariant matrix insertions (Corollary 5.21) using the corner value and one radial coefficient of each matrix. For insertions of the form $Q_{\mathrm{RR}}g$ or $Q_{\mathrm{AA}}g$ with invariant g, the semigroup preserves the corner value and evolves the radial coefficient by the radial semigroup. The full matrix decomposition also contains an independent scalar function multiplying the identity, as in Lemma 6.12. Throughout this section, $X > 0$, $z > 0$, and $\delta \in (0, 1/2)$ are fixed. We write $\mathsf{G}_{X,z} = X\mathsf{C}_1 + z\, d_1(Q)$, with $d_1(Q) = \frac{1}{2}\,\mathrm{STr}(\Lambda Q)$, and $U_t = e^{-t\mathsf{G}_{X,z}}$ for $t \ge 0$, realized as in Proposition 5.15. The pairing $\mathcal{B}_1$ is the invariant bilinear Berezin pairing of Definition 5.6, and $\int A = \mathcal{B}_1(1, A)$.

Recall from (5.1) that the elements Y of $\mathfrak{k}$ are block diagonal in the flavors, with the flavor blocks Y_{RR} and Y_{AA}, and recall the advanced and retarded subalgebras $\mathfrak{k}_{\mathrm{A}}$ and $\mathfrak{k}_{\mathrm{R}}$ of Lemma 5.18, which consist of the elements with $Y_{\mathrm{RR}} = 0$, respectively $Y_{\mathrm{AA}} = 0$. Recall the derivations $\widehat{Y}$ of Lemma 5.4 and the auxiliary parameters ϑ_Y used there. We call a superfunction invariant if it is invariant under $\mathfrak{k}$ in the sense of Section 5.2. Recall from the paragraph before Corollary 5.21 the advanced-invariant (retarded-invariant)

superfunctions, which are invariant under $\mathfrak{k}_{\mathrm{A}}$ (under $\mathfrak{k}_{\mathrm{R}}$), and the retarded-covariant and advanced-covariant matrices of superfunctions. At rank one, these matrices are $(1|1)$ supermatrices of superfunctions, which we write in bold. A retarded-covariant matrix is advanced-invariant entrywise, since $Y_{\mathrm{RR}} = 0$ for $Y \in \mathfrak{k}_{\mathrm{A}}$. By (5.6), the retarded block Q_{RR} of Q is retarded-covariant, and Q_{AA} is advanced-covariant.

We use the following facts from the preceding sections.

(i) *Localization.* If A is an advanced-invariant or retarded-invariant smooth superfunction with polynomially integrable coefficients, then $\int A = A_{\varnothing}(\Lambda)$, by Lemmas 5.18 and 5.19, as in the proof of Corollary 5.21. The same holds entrywise for matrices of superfunctions.

(ii) *Polar coordinates.* We use the entries of Q in polar coordinates of Lemma 6.1(b), the radial representatives of Lemma 6.7(a), and the top coefficients $[A]_{\mathrm{top}}$ and $\langle A \rangle_{\mathrm{top}}$ of Lemma 6.1(c).

(iii) *Radial reduction.* The semigroup U_t preserves invariant superfunctions and acts on them as in Lemma 6.7(c). The radial representative of $U_t 1$ is $\Psi_t^{X,z}$, by Lemma 6.7(d), and its value at the corner is one, by (6.14).

(iv) *Doob transform.* We use the identity (6.19), which the proof of Lemma 6.7 derives from Lemma 6.5.

All superfunctions on $\mathcal{M}_1$ below, and all entries of matrices of superfunctions below, are built from finitely many polynomials in the entries of Q and semigroup factors U_t applied to 1, in which the outermost semigroup factors have times at least δ. By Proposition 5.15(b),(d) and Lemma 5.14, such superfunctions are smooth and rapidly decreasing, and by the Cauchy–Schwarz inequality their products have polynomially integrable coefficients in the sense of Lemma 5.19. We use this without further mention.

Let $\mathbf{M}$ be a retarded-covariant matrix of superfunctions as above. Its entries are advanced-invariant, and the case $k = 1$ of Corollary 5.21, with $F = \mathbf{M}$, shows that $\mathbf{M}_{\varnothing}(\Lambda) = c_\Lambda \mathrm{Id}$ for a scalar c_Λ and that $\mathcal{B}_1(f, \mathbf{M}) = f_{\varnothing}(\Lambda)\, c_\Lambda \mathrm{Id}$ for every invariant superfunction f as above. The same holds with the roles of the flavors exchanged. The following corollary adds that the semigroup preserves the covariance and the value at Λ.

Corollary 6.11. *Let* $\mathbf{M}$ *be a retarded-covariant matrix of superfunctions as above, and let* c_Λ *be the scalar with* $\mathbf{M}_{\varnothing}(\Lambda) = c_\Lambda \mathrm{Id}$. *Then, for all* $t \geq 0$, *the matrix* $U_t\mathbf{M}$ *is retarded-covariant, and*

$$(U_t\mathbf{M})_{\varnothing}(\Lambda) = c_\Lambda \mathrm{Id}. \tag{6.22}$$

The same holds with the roles of the retarded and the advanced flavor exchanged.

Proof. We first show that $U_t\mathbf{M}$ is retarded-covariant. Let $Y \in \mathfrak{k}$ be homogeneous, and set

$$\Phi_1(t) = \vartheta_Y \widehat{Y} U_t \mathbf{M} - [\vartheta_Y Y_{\mathrm{RR}}, U_t \mathbf{M}].$$

The derivation $\widehat{Y}$ commutes with $\mathsf{G}_{X,z}$, by Lemma 5.8 and the invariance of d_1, and the commutator with the constant matrix $\vartheta_Y Y_{\mathrm{RR}}$ commutes with $\mathsf{G}_{X,z}$ entrywise. Since $\mathbf{M}$ is rapidly decreasing, (5.25) shows that the entries of $\mathbf{M}$ lie in $D(\mathsf{G}^2_{(r)})$ for all weights $r \in \mathbb{R}$. Recall from the first step of the proof of Corollary 5.20 that, for such initial data, $t \mapsto \widehat{Y} U_t \mathbf{M}$ is continuously differentiable, so that Φ_1 is a strong solution of $\Phi_1' = -\mathsf{G}_{X,z}\Phi_1$ with $\Phi_1(0) = 0$, and that such a solution vanishes, by (5.31). Hence $\Phi_1 = 0$.

We now prove (6.22). Let $g = U_\delta 1$. By fact (iii), g is invariant, and $U_t g = U_{t+\delta} 1$ has the value one at Λ. Since $U_t \mathbf{M}$ is retarded-covariant, Corollary 5.21 implies that $(U_t\mathbf{M})_\varnothing(\Lambda) = c_\Lambda(t)\mathrm{Id}$ for a scalar $c_\Lambda(t)$. We apply Corollary 5.21 twice, together with the symmetry of $\mathcal{B}_1$ under U_t of Proposition 5.15(c), and we obtain

$$c_\Lambda(t)\mathrm{Id} = g_\varnothing(\Lambda)\, c_\Lambda(t)\mathrm{Id} = \mathcal{B}_1(g, U_t\mathbf{M}) = \mathcal{B}_1(U_t g, \mathbf{M}) = (U_t g)_\varnothing(\Lambda)\, c_\Lambda \mathrm{Id} = c_\Lambda \mathrm{Id}.$$

The advanced case is the same. □

Localization evaluates every pairing whose integrand is invariant under the stabilizer of one flavor, and this covers most of the pairings that we need. The cross pairings, which have a retarded and an advanced insertion, mix the two flavors, and so does the contact multiplier, whose pairing Lemma 11.2 evaluates in Section 11. They require the polar chart near the corner, where it degenerates, and the corner term of the radial integration formula Lemma 6.3 is the boundary term that they produce. The main step of this section is the evaluation of these mixed pairings. Such terms are an instance of the boundary terms of Berezin integration on noncompact supermanifolds, first described by Rothstein [110], which are known in physics as Efetov–Wegner terms; see [139, Section II.A] and [76].

We now describe the evolution of advanced-covariant matrices. The following lemma decomposes such a matrix as $a\mathrm{Id} + bQ_{\mathrm{AA}}$ and evaluates the mixed pairings below using only its corner value and b. For a matrix $Q_{\mathrm{AA}}g$ with invariant g, U_t preserves the corner value and evolves b by the radial diffusion. For the infinite wire, the corresponding structure in the physics literature is the flat bilinear form of [71, (16)–(17)].

Lemma 6.12. *Let* $\mathbf{M}$ *be an advanced-covariant matrix of superfunctions as above. Then there exist functions* a *and* b *of* (B, F) *on* $D \setminus \{(1,1)\}$ *such that, in the polar chart,*

$$\mathbf{M} = a\mathrm{Id} + b\, Q_{\mathrm{AA}}, \qquad b = -\frac{\mathrm{STr}\,\mathbf{M}}{d}. \tag{6.23}$$

Let g *be an invariant superfunction as above, and let* $\mathbf{M} = Q_{\mathrm{AA}}g$*, so that* $b = g$*. Then for all* $t \geq 0$*, the coefficient* b_t *in the decomposition* (6.23) *of* $U_t\mathbf{M}$ *and the value of the body of* $U_t\mathbf{M}$ *at* Λ *are*

$$b_t = e^{t(X\Delta_D - zd)} g, \qquad (U_t\mathbf{M})_\varnothing(\Lambda) = -g_\varnothing(\Lambda)\mathrm{Id}. \tag{6.24}$$

Moreover, for every invariant superfunction f as above,

$$\mathcal{B}_1\big(f, Q_{\mathrm{RR}}^{\mathrm{FF}}(U_t\mathbf{M})^{\mathrm{FF}}\big) = -f_\varnothing(\Lambda)\, g_\varnothing(\Lambda) - \int_D f\, e^{t(X\Delta_D - zd)} g \,\mathrm{d}B\,\mathrm{d}F. \tag{6.25}$$

The same statements hold with the roles of the two flavors exchanged, for $\mathbf{M} = Q_{\mathrm{RR}}g$, with $b = \mathrm{STr}\,\mathbf{M}/d$ and $(U_t\mathbf{M})_\varnothing(\Lambda) = g_\varnothing(\Lambda)\mathrm{Id}$.

Proof. We first prove (6.23). By (6.2), a point with $d > 0$ has the form $Q = e^{-\Theta_\xi} Q_0 e^{\Theta_\xi}$, where Q_0 has the same body and vanishing odd coordinates, and Θ_ξ is a linear combination of the ξ_a with coefficients in the odd part of $\mathfrak{k}$. The entries of $\mathbf{M}(Q_0)$ are ordinary functions, so the odd entries vanish, and $\mathbf{M}(Q_0) = \mathrm{diag}(f^{\mathrm{B}}, f^{\mathrm{F}})$. The phases are rotated by even elements of $\mathfrak{k}$ that act on $\mathbf{M}(Q_0)$ by conjugation with diagonal unitary matrices, so the integrated form of covariance below shows that f^{B} and f^{F} are functions of (B, F). Since $Q_{\mathrm{AA}}(Q_0) = \mathrm{diag}(-B, -F)$ and $d > 0$, we have $\mathbf{M}(Q_0) = a\mathrm{Id} + b\, Q_{\mathrm{AA}}(Q_0)$ with

$$b = -\frac{f^{\mathrm{B}} - f^{\mathrm{F}}}{d}, \qquad a = f^{\mathrm{B}} + bB.$$

Let $k_{\mathrm{A}}(r)$ be the advanced block of $e^{-r\Theta_\xi}$, and set

$$\mathbf{M}(r) = k_{\mathrm{A}}(r)^{-1}\, \mathbf{M}\big(e^{-r\Theta_\xi} Q_0 e^{r\Theta_\xi}\big)\, k_{\mathrm{A}}(r), \qquad r \in [0,1].$$

Since Θ_ξ is a combination of odd elements of $\mathfrak{k}$, the chain rule and covariance imply that $\partial_r \mathbf{M}(r) = 0$, as in the proof of Lemma 6.7(a). Therefore

$$\mathbf{M}(Q) = k_{\mathrm{A}}(1)\mathbf{M}(Q_0)k_{\mathrm{A}}(1)^{-1} = a\mathrm{Id} + b\, Q_{\mathrm{AA}}(Q).$$

Taking supertraces and using $\mathrm{STr}\, Q_{\mathrm{AA}} = -d$, which follows from Lemma 6.1(b), we find $\mathrm{STr}\,\mathbf{M} = -bd$.

Next, let $\mathbf{M} = Q_{\mathrm{AA}}g$, so that $\mathrm{STr}\,\mathbf{M} = -dg$. The supertrace is a linear combination of entries, and U_t acts on matrices of superfunctions entrywise, so $\mathrm{STr}\, U_t\mathbf{M} = U_t\,\mathrm{STr}\,\mathbf{M}$. The superfunction $\mathrm{STr}\,\mathbf{M}$ is invariant, and g is bounded. By (iii) and (6.19) with $\varphi = g$, we obtain

$$U_t(dg) = d\, e^{t(X\Delta_D - zd)} g$$

on D°, and the first identity in (6.24) follows from (6.23) for $U_t\mathbf{M}$. Further, $\mathbf{M}_\varnothing(\Lambda) = -g_\varnothing(\Lambda)\mathrm{Id}$, since $(Q_{\mathrm{AA}})_\varnothing(\Lambda) = -\mathrm{Id}$, and Corollary 6.11, with the flavors exchanged, implies the second identity in (6.24).

We now prove (6.25), and we apply Lemma 6.3 to $A = f Q_{\mathrm{RR}}^{\mathrm{FF}}(U_t\mathbf{M})^{\mathrm{FF}}$. Since $(Q_{\mathrm{RR}}^{\mathrm{FF}})_\varnothing(\Lambda) = 1$, we have $A_\varnothing(\Lambda) = -f_\varnothing(\Lambda)\, g_\varnothing(\Lambda)$. By Lemma 6.1(b) and (6.23),

$$A = f\,(F + d\, n_{\mathrm{R}})\big(a_t - b_t F - b_t d\, n_{\mathrm{A}}\big), \qquad [A]_{\mathrm{top}} = -d^2 f b_t.$$

In particular, $d^{-2}[A]_{\mathrm{top}} = -fb_t$ is independent of the phases, so that $\langle A\rangle_{\mathrm{top}} = [A]_{\mathrm{top}}$. It is integrable on D, since $|f| \le C(1+B)^{-3}$ by Lemma 6.6(c) (the outermost times are at least δ, and $|\Psi_t^{X,z}| \le \Psi_t^{X,\mathrm{Re}\, z}$), and $|b_t| \le \sup|g|$. Applying Lemma 6.3, we conclude that (6.25) holds. With the flavors exchanged, we have $\mathrm{STr}(Q_{\mathrm{RR}}g) = dg$. Again $[Q_{\mathrm{AA}}^{\mathrm{FF}}(a_t + b_t Q_{\mathrm{RR}}^{\mathrm{FF}})]_{\mathrm{top}} = -d^2 b_t$, and the body of this product at Λ is $-g_\varnothing(\Lambda)$. □

For an advanced-covariant matrix $\mathbf{M}$ of superfunctions as above, we call the matrix $\mathbf{M}_\varnothing(\Lambda) = c_\Lambda \mathrm{Id}$ of Corollary 5.21 its *corner channel*, and the coefficient b of Q_{AA} in (6.23) its *radial channel*; for retarded-covariant matrices, the coefficient of Q_{RR} plays the role of b. By Lemma 6.12, the semigroup U_t preserves the corner channel $-g_\varnothing(\Lambda)\mathrm{Id}$ of $Q_{\mathrm{AA}}g$ and acts on its radial channel g by the radial semigroup. In the pairing (6.25), which mixes the two flavors, the corner channel contributes through the corner term of Lemma 6.3, and the radial channel through the bulk term.

We now evaluate the pairings of two singlet states with one or two diagonal entries of Q inserted. In Part III, these are the limits that Theorem 7.6(b) identifies for the multipliers of single resolvent entries (Sections 8.6 and 11.2).

Lemma 6.13. *Let $\delta \le x \le y \le 1-\delta$, and set $t = y - x$. Then, with $\Psi_x = \Psi_x^{X,z}$ and $\Psi_{1-y} = \Psi_{1-y}^{X,z}$,*

$$\mathcal{B}_1\big(\Psi_x, Q_{\mathrm{RR}}^{\mathrm{FF}} U_t Q_{\mathrm{AA}}^{\mathrm{FF}} \Psi_{1-y}\big) = \mathcal{B}_1\big(\Psi_x, Q_{\mathrm{AA}}^{\mathrm{FF}} U_t Q_{\mathrm{RR}}^{\mathrm{FF}} \Psi_{1-y}\big) = -1 - K_X(x,y;z), \tag{6.26}$$

and

$$\mathcal{B}_1\big(\Psi_x, U_t \Psi_{1-y}\big) = 1, \qquad \mathcal{B}_1\big(\Psi_x, Q_{\mathrm{RR}}^{\mathrm{FF}} U_t \Psi_{1-y}\big) = 1, \qquad \mathcal{B}_1\big(\Psi_x, U_t Q_{\mathrm{AA}}^{\mathrm{FF}} \Psi_{1-y}\big) = -1. \tag{6.27}$$

Further, for $c, c' \in \{\mathrm{BB}, \mathrm{FF}\}$,

$$\begin{gathered}\mathcal{B}_1\big(\Psi_x, Q_{\mathrm{RR}}^{c} U_t \Psi_{1-y}\big) = \mathcal{B}_1\big(\Psi_x, U_t Q_{\mathrm{RR}}^{c} \Psi_{1-y}\big) = 1, \qquad \mathcal{B}_1\big(\Psi_x, U_t Q_{\mathrm{AA}}^{c} \Psi_{1-y}\big) = -1, \\ \mathcal{B}_1\big(\Psi_x, Q_{\mathrm{RR}}^{c} U_t Q_{\mathrm{RR}}^{c'} \Psi_{1-y}\big) = \mathcal{B}_1\big(\Psi_x, Q_{\mathrm{AA}}^{c} U_t Q_{\mathrm{AA}}^{c'} \Psi_{1-y}\big) = 1.\end{gathered} \tag{6.28}$$

These identities include $t = 0$, where $U_0 = \mathrm{Id}$. Here δ enters only through the outermost times x and $1-y$, so the lemma applies to all $0 < x \le y < 1$ with $\delta = \min(x, 1-y)$.

Proof. By (iii), $\Psi_x = U_x 1$ and $\Psi_{1-y} = U_{1-y}1$ are invariant superfunctions with the value one at Λ. The first identity in (6.26) is (6.25) with $f = \Psi_x$ and $g = \Psi_{1-y}$, together with the definition (1.6) of K_X. The second is the version of (6.25) with the flavors exchanged.

In (6.27), the first integrand is invariant, and the second is advanced-invariant, since $U_t\Psi_{1-y} = \Psi_{1-x}$ is invariant and $Q_{\mathrm{RR}}^{\mathrm{FF}}$ is an entry of Q_{RR}. Using (i), we evaluate these pairings at Λ, where the bodies of the integrands take the values 1 and $(Q_{\mathrm{RR}}^{\mathrm{FF}})_\varnothing(\Lambda) = 1$. The third integrand is Ψ_x times an entry of the retarded-invariant matrix $U_t(Q_{\mathrm{AA}}\Psi_{1-y})$, whose body at Λ is $-\mathrm{Id}$ by Lemma 6.12. The identities (6.28) follow in the same way.

The superfunctions Q^c_{RR} are entries of Q_{RR}, and $U_t(Q_{\mathrm{RR}}\Psi_{1-y})$ is retarded-covariant and advanced-invariant with body Id at Λ, by Corollary 6.11. Every integrand in (6.28) with retarded insertions is therefore advanced-invariant, and by (i) its integral is the value of its body at Λ, which is one. With the flavors exchanged, $U_t(Q_{\mathrm{AA}}\Psi_{1-y})$ has the body $-\mathrm{Id}$ at Λ, and $(Q^c_{\mathrm{AA}})_\varnothing(\Lambda) = -1$, and the remaining values follow. □

6.5 How the kernels arise

We now compute the first two derivatives of the partition function of rank one in its source. This is where the singlet state and the kernels of Definition 1.4 meet the partition function. The outer semigroup factors become singlet states, the pairings with one retarded and one advanced insertion produce K_X, and the integral over the two positions along the wire produces I_X.

We first expand the semigroup in the sources by a Dyson series and show that the partition function is entire in the sources. In the following lemma only, the rank $k \geq 1$ is arbitrary. Section 8.4 uses it at every rank, and the proof of Lemma 6.15 below at rank one.

Lemma 6.14. *Let $k \geq 1$, $X > 0$, and $\varepsilon > 0$, let $z = 2\pi\varepsilon$, and let $U_s = e^{-s\mathsf{G}_{X,z}}$ for $s \geq 0$. Let $q > q_1/2$ with $4q \in \mathbb{Z}$, where q_1 is the exponent of Lemma 5.1(c). For $n \geq 0$, let*

$$\Delta_n = \Big\{ s \in [0,1]^{n+1} : \sum_{i=0}^{n} s_i = 1 \Big\},$$

with the Lebesgue measure $\mathrm{d}s = \mathrm{d}s_0 \cdots \mathrm{d}s_{n-1}$, of total mass $1/n!$.

(a) There exists $C \geq 1$ such that

$$\|U_s\|_{L^2_{\pm q} \to L^2_{\pm q}} \leq C, \quad 0 \leq s \leq 1, \qquad \|U_s\|_{L^2_{-q} \to L^2_q} \leq C s^{-2q}, \quad 0 < s \leq 1. \tag{6.29}$$

(b) Let V be an operator that is bounded on L^2_q and on L^2_{-q}, and let $\|V\|_{\pm q}$ denote the larger of its two operator norms. Then, for all $n \geq 0$ and $s \in \Delta_n$,

$$\big\| U_{s_0} V U_{s_1} V \cdots V U_{s_n} \big\|_{L^2_{-q} \to L^2_q} \leq C^{n+1} (n+1)^{2q} \|V\|^n_{\pm q}. \tag{6.30}$$

(c) For a complex diagonal $2k \times 2k$ matrix Γ_{F}, let $\Gamma = \mathrm{diag}(\varepsilon\Lambda_{\mathrm{B}} | \Gamma_{\mathrm{F}})$, and set

$$V_\Gamma = -\pi \operatorname{Tr}\big(Q^{\mathrm{FF}}(\Gamma_{\mathrm{F}} - \varepsilon\Lambda_{\mathrm{F}})\big), \qquad \mathcal{S}_\Gamma = \sum_{n=0}^{\infty} (-1)^n \int_{\Delta_n} U_{s_0} V_\Gamma U_{s_1} V_\Gamma \cdots V_\Gamma U_{s_n} \, \mathrm{d}s. \tag{6.31}$$

The series converges in $L(L^2_{-q}, L^2_q)$, locally uniformly in the diagonal entries of Γ_{F}, and $\mathcal{S}_\Gamma = e^{-\mathsf{G}_{X,\Gamma}}$ on L^2_{-q}, with the semigroup of Proposition 5.15. The function $\Gamma_{\mathrm{F}} \mapsto \int e^{-\mathsf{G}_{X,\Gamma}} 1 =$

$\mathcal{B}_k(1, \mathcal{S}_\Gamma 1)$ is entire in the diagonal entries of Γ_F. In particular, the partition function $\mathcal{Z}^{(k)}_{X,\varepsilon}(\mathbf{u}, \mathbf{b})$ of Definition 5.16, as a function of the diagonal entries of Γ_F, is the restriction of an entire function on $\mathbb{C}^{2k}$.

Proof. We begin with (a). By Proposition 5.15(e), the semigroup U_s is the one of Proposition 5.15 with $\Gamma = \varepsilon\Lambda$. Applying Proposition 5.15(b) on $L^2_{\pm q}$ with $m = 0$, and from L^2_{-q} with $m = 4q$, we obtain (a), since $U_0 = \mathrm{Id}$.

For (b), let $s \in \Delta_n$. Some s_i is at least $1/(n+1)$. We bound the factor U_{s_i} from L^2_{-q} to L^2_q by (6.29), which gives the factor $C(n+1)^{2q}$, the factors to its right on L^2_{-q}, and the factors to its left on L^2_q. This proves (6.30).

We now prove (c). By Lemma 5.14(a), the entries of Q^{FF} are multipliers of degree zero. Hence there exists C_q such that $\|V_\Gamma\|_{\pm q} \le C_q \|\Gamma_F - \varepsilon\Lambda_F\|_{\mathrm{op}}$ for all Γ_F. By (6.30) with $V = V_\Gamma$, and since Δ_n has the mass $1/n!$, the nth term of the series in (6.31) has norm at most

$$\frac{C^{n+1}(n+1)^{2q}}{n!}\big(C_q\|\Gamma_F - \varepsilon\Lambda_F\|_{\mathrm{op}}\big)^n$$

in $L(L^2_{-q}, L^2_q)$. The nth term is a homogeneous polynomial of degree n in the diagonal entries of $\Gamma_F - \varepsilon\Lambda_F$, with values in $L(L^2_{-q}, L^2_q)$. Therefore the series converges locally uniformly, and $\Gamma_F \mapsto \mathcal{S}_\Gamma$ is entire.

Next, we identify $\mathcal{S}_\Gamma$ with the semigroup. By (5.22), we have $\mathsf{G}_{X,\Gamma} = \mathsf{G}_{X,z} + V_\Gamma$ on the compactly supported smooth superfunctions. Since V_Γ is bounded, the closure of $\mathsf{G}_{X,\Gamma}$ in L^2_{-q} of Proposition 5.15(a) is the closure of $\mathsf{G}_{X,z}$ plus V_Γ, with the same domain; here we use one norm on L^2_{-q} for both realizations, which is possible since the realizations do not depend on ω and the norms for different ω are equivalent. The semigroup of a bounded perturbation of a generator is given by the Dyson series [72, Chapter IX], which converges in $L(L^2_{-q})$ and is the series in (6.31). Since L^2_q embeds continuously in L^2_{-q}, we obtain $e^{-\mathsf{G}_{X,\Gamma}} = \mathcal{S}_\Gamma$ on L^2_{-q}.

Finally, $1 \in L^2_{-q}$ by Proposition 5.15(d), and $\mathcal{B}_k$ is bounded on $L^2_{-q} \times L^2_q$ by Proposition 5.15(c). Hence $\int e^{-\mathsf{G}_{X,\Gamma}} 1 = \mathcal{B}_k(1, e^{-\mathsf{G}_{X,\Gamma}}1) = \mathcal{B}_k(1, \mathcal{S}_\Gamma 1)$ is entire in the diagonal entries of Γ_F. For the sources $(\mathbf{u}, \mathbf{b})$ of Definition 5.16, this is $\mathcal{Z}^{(k)}_{X,\varepsilon}(\mathbf{u}, \mathbf{b})$, by Definition 5.16 and Proposition 5.15(d), and the map from $(\mathbf{u}, \mathbf{b})$ to the diagonal entries of Γ_F is injective. □

The following lemma is the computation in the sigma model that the proof of Theorem 1.6(c) in Section 8.6 uses.

Lemma 6.15. *Let $X > 0$ and $\varepsilon > 0$, let $z = 2\pi\varepsilon$ and $U_s = e^{-s\mathsf{G}_{X,z}}$ for $s \ge 0$, and let*

$$\mathcal{Z}(t) = \mathcal{Z}^{(1)}_{X,\varepsilon}(0, \varepsilon + t), \quad t > -\varepsilon, \qquad m_F = \mathrm{Tr}(\Lambda_F Q^{\mathrm{FF}}) = Q^{\mathrm{FF}}_{\mathrm{RR}} - Q^{\mathrm{FF}}_{\mathrm{AA}}.$$

Then $\mathcal{Z}$ is smooth on $(-\varepsilon,\infty)$, and

$$\begin{aligned}\mathcal{Z}'(0) &= \pi\int_0^1 \mathcal{B}_1\big(1, U_s m_{\mathrm{F}} U_{1-s}1\big)\,\mathrm{d}s,\\ \mathcal{Z}''(0) &= 2\pi^2\int_{0<x<y<1} \mathcal{B}_1\big(1, U_x m_{\mathrm{F}} U_{y-x} m_{\mathrm{F}} U_{1-y}1\big)\,\mathrm{d}x\,\mathrm{d}y.\end{aligned} \tag{6.32}$$

For $0 < x \le y < 1$, we have

$$\mathcal{B}_1\big(1, U_x m_{\mathrm{F}} U_{1-x}1\big) = 2, \qquad \mathcal{B}_1\big(1, U_x m_{\mathrm{F}} U_{y-x} m_{\mathrm{F}} U_{1-y}1\big) = 4 + 2K_X(x,y;z). \tag{6.33}$$

In particular,

$$\mathcal{Z}'(0) = 2\pi, \qquad \mathcal{Z}''(0) = 4\pi^2\big(1 + I_X(2\pi\varepsilon)\big). \tag{6.34}$$

Proof. We begin with the generator. At rank one, $u_1 = 0$ and $b_1 = \varepsilon + t$, so that $\Gamma_{\mathrm{F}} = (\varepsilon+t)\Lambda_{\mathrm{F}}$ and $\Gamma_{\mathrm{F}} - \varepsilon\Lambda_{\mathrm{F}} = t\Lambda_{\mathrm{F}}$. Inserting this into (5.22), we obtain

$$\mathsf{G}_{X,\Gamma(0,\varepsilon+t)} = \mathsf{G}_{X,z} - \pi t\, m_{\mathrm{F}}. \tag{6.35}$$

We now expand the semigroup in t. We apply Lemma 6.14 with $k = 1$ and $\Gamma_{\mathrm{F}} = (\varepsilon+t)\Lambda_{\mathrm{F}}$, for which $V_\Gamma = -\pi t\, m_{\mathrm{F}}$, by (6.35), and we let q, C, and Δ_n be as there. The series (6.31) becomes

$$e^{-\mathsf{G}_{X,\Gamma(0,\varepsilon+t)}} = \sum_{n=0}^{\infty} (\pi t)^n \int_{\Delta_n} U_{s_0} m_{\mathrm{F}} U_{s_1} m_{\mathrm{F}} \cdots m_{\mathrm{F}} U_{s_n}\,\mathrm{d}s, \tag{6.36}$$

with convergence in $L(L^2_{-q}, L^2_q)$, and Lemma 6.14(c) gives

$$\mathcal{Z}(t) = \int e^{-\mathsf{G}_{X,\Gamma(0,\varepsilon+t)}}1 = \mathcal{B}_1\big(1, e^{-\mathsf{G}_{X,\Gamma(0,\varepsilon+t)}}1\big).$$

By Lemma 5.14(a), the entries of Q^{FF} are multipliers of degree zero, so that multiplication by m_{F} is bounded on L^2_q and on L^2_{-q}, and we let C_q bound both norms. Inserting (6.36), and using Lemma 6.14(b) with $V = m_{\mathrm{F}}$ and the boundedness of $\mathcal{B}_1$ on $L^2_{-q} \times L^2_q$ of Proposition 5.15(c), we find that, for all $t > -\varepsilon$,

$$\begin{gathered}\mathcal{Z}(t) = \sum_{n=0}^{\infty} (\pi t)^n c_n, \qquad c_n = \int_{\Delta_n} \mathcal{B}_1\big(1, U_{s_0} m_{\mathrm{F}} U_{s_1} \cdots m_{\mathrm{F}} U_{s_n}1\big)\,\mathrm{d}s,\\ |c_n| \le \frac{C' C^{n+1} (n+1)^{2q} C_q^n}{n!},\end{gathered} \tag{6.37}$$

where C' depends only on q. The power series converges for all t, so that $\mathcal{Z}$ is smooth on $(-\varepsilon,\infty)$, with $\mathcal{Z}'(0) = \pi c_1$ and $\mathcal{Z}''(0) = 2\pi^2 c_2$. In the coordinates $x = s_0$ and $y = s_0 + s_1$, the simplex Δ_2 becomes $\{0 < x < y < 1\}$, with $\mathrm{d}s = \mathrm{d}x\,\mathrm{d}y$, and we obtain (6.32).

We now evaluate the pairings in (6.32) with the rank-one calculus of Section 6.4, applied with the present z. Let $0 < x \le y < 1$, and write $\Psi_s = \Psi_s^{X,z}$. By fact (iii) of Section 6.4, the superfunctions $U_x 1$ and $U_{1-y}1$ are invariant, with the radial representatives Ψ_x and Ψ_{1-y}, and we denote them by the same letters. Their values at Λ are one, by (6.14). The transpose symmetry of Proposition 5.15(c) implies that $\mathcal{B}_1(1, U_x f) = \mathcal{B}_1(\Psi_x, f)$. The pairings in (6.32) are therefore those of Lemma 6.13, with $\delta = \min(x, 1-y)$ and with the time $y - x$ in place of the t of that lemma.

For the first pairing, we expand m_{F} and use the second and the third identities in (6.27) with $y = x$. We obtain

$$\mathcal{B}_1\big(1, U_x m_{\mathrm{F}} U_{1-x} 1\big) = \mathcal{B}_1\big(\Psi_x, Q_{\mathrm{RR}}^{\mathrm{FF}} \Psi_{1-x}\big) - \mathcal{B}_1\big(\Psi_x, Q_{\mathrm{AA}}^{\mathrm{FF}} \Psi_{1-x}\big) = 1 - (-1) = 2.$$

For the second pairing, we expand both factors m_{F}, and we obtain four pairings. The two with insertions of one flavor equal one, by (6.28) with $c = c' = \mathrm{FF}$, and the two with insertions of both flavors equal $-1 - K_X(x, y; z)$, by (6.26). In the last two, the corner channel $-\mathrm{Id}$ of the advanced-covariant matrix $U_{y-x}(Q_{\mathrm{AA}} \Psi_{1-y})$, which the semigroup preserves, contributes the corner term -1 of the radial integration formula of Lemma 6.3. Its radial channel $e^{(y-x)(X\Delta_D - zd)} \Psi_{1-y}$, which evolves by the radial semigroup, contributes the bulk term $-K_X(x, y; z)$, by Lemma 6.12. Since $m_{\mathrm{F}} U_{y-x} m_{\mathrm{F}}$ contains the two pairings of one flavor with the sign $+$ and the two others with the sign $-$, we conclude that

$$\mathcal{B}_1\big(1, U_x m_{\mathrm{F}} U_{y-x} m_{\mathrm{F}} U_{1-y} 1\big) = 1 + 1 + 2\big(1 + K_X(x, y; z)\big) = 4 + 2K_X(x, y; z),$$

which completes the proof of (6.33).

It remains to integrate. The integral of K_X over $\{0 < x < y < 1\}$ converges absolutely, by Lemma 6.6(d), and the definition (1.8) shows that it equals $I_X(z)$. Inserting (6.33) into (6.32), we find

$$\mathcal{Z}'(0) = \pi \int_0^1 2 \, \mathrm{d}s = 2\pi,$$
$$\mathcal{Z}''(0) = 2\pi^2 \int_{0<x<y<1} \big(4 + 2K_X(x, y; z)\big) \, \mathrm{d}x \, \mathrm{d}y = 2\pi^2 \big(2 + 2I_X(z)\big),$$

which is (6.34), since $z = 2\pi\varepsilon$. □

The first identity in (6.34) expresses the unit intensity of the limiting point process, and the second determines its pair correlation through Remark 8.9; they are the two limits in (1.10). The contact kernel $\widetilde{K}_X$ arises in the same way from the product $Q_{\mathrm{RR}}^{\mathrm{BB}} Q_{\mathrm{AA}}^{\mathrm{BB}} + Q_{\mathrm{RA}}^{\mathrm{BB}} Q_{\mathrm{AR}}^{\mathrm{BB}}$ of entries of Q at a single site. By Lemma 6.1(c), we have

$$\big\langle Q_{\mathrm{RR}}^{\mathrm{BB}} Q_{\mathrm{AA}}^{\mathrm{BB}} + Q_{\mathrm{RA}}^{\mathrm{BB}} Q_{\mathrm{AR}}^{\mathrm{BB}} \big\rangle_{\mathrm{top}} = -d^2 - d(B + F) = -2Bd,$$

and after multiplication by the density d^{-2} of Lemma 6.3, it produces the weight B/d of (1.7). The corresponding pairing is evaluated in Lemma 11.2.

References

[1] Shmuel Agmon. *Lectures on Exponential Decay of Solutions of Second-Order Elliptic Equations*. Princeton University Press, Princeton, 1982.

[2] Alexander Alldridge and Joachim Hilgert. Invariant Berezin integration on homogeneous supermanifolds. *J. Lie Theory*, 20(1):65–91, 2010.

[3] Alexander Alldridge, Joachim Hilgert, and Wolfgang Palzer. Berezin integration on non-compact supermanifolds. *J. Geom. Phys.*, 62(2):427–448, 2012. doi: 10.1016/j.geomphys.2011.11.005.

[4] A. Altland and D. Fuchs. Spectral statistics of mesoscopic wires: crossover from Wigner–Dyson to Poisson regime. *Phys. Rev. Lett.*, 74(21):4269–4272, 1995. doi: 10.1103/PhysRevLett.74.4269.

[5] B. L. Al'tshuler and B. I. Shklovskiĭ. Repulsion of energy levels and conductivity of small metal samples. *Sov. Phys. JETP*, 64(1):127–135, 1986. Russian original: Zh. Eksp. Teor. Fiz. **91** (1986), 220–234.

[6] Zhigang Bao and László Erdős. Delocalization for a class of random block band matrices. *Probab. Theory Related Fields*, 167(3-4):673–776, 2017. doi: 10.1007/s00440-015-0692-y.

[7] Roland Bauerschmidt, David C. Brydges, and Gordon Slade. *Introduction to a Renormalisation Group Method*, volume 2242 of *Lecture Notes in Mathematics*. Springer, Singapore, 2019. doi: 10.1007/978-981-32-9593-3.

[8] Felix Alexandrovich Berezin. *Introduction to Superanalysis*. D. Reidel (Springer Netherlands), 1987. doi: 10.1007/978-94-017-1963-6.

[9] Ya. M. Blanter and A. D. Mirlin. Correlations of eigenfunctions in disordered systems. *Phys. Rev. E*, 55(6):6514–6518, 1997. doi: 10.1103/PhysRevE.55.6514.

[10] Paul Bourgade. Random band matrices. In *Proceedings of the International Congress of Mathematicians (ICM 2018)*, pages 2759–2783. World Scientific, 2019. doi: 10.1142/9789813272880_0159.

[11] Paul Bourgade, László Erdős, Horng-Tzer Yau, and Jun Yin. Universality for a class of random band matrices. *Adv. Theor. Math. Phys.*, 21(3):739–800, 2017. doi: 10.4310/ATMP.2017.v21.n3.a5.

[12] Paul Bourgade, Fan Yang, Horng-Tzer Yau, and Jun Yin. Random band matrices in the delocalized phase, II: Generalized resolvent estimates. *J. Stat. Phys.*, 174(6):

1189–1221, 2019. doi: 10.1007/s10955-019-02229-z.

[13] Paul Bourgade, Horng-Tzer Yau, and Jun Yin. Random band matrices in the delocalized phase I: Quantum unique ergodicity and universality. *Comm. Pure Appl. Math.*, 73(7):1526–1596, 2020. doi: 10.1002/cpa.21895.

[14] Massimo Campanino and Abel Klein. A supersymmetric transfer matrix and differentiability of the density of states in the one-dimensional Anderson model. *Comm. Math. Phys.*, 104(2):227–241, 1986.

[15] Giulio Casati, Luca Molinari, and Felix Izrailev. Scaling properties of band random matrices. *Phys. Rev. Lett.*, 64(16):1851–1854, 1990. doi: 10.1103/PhysRevLett.64.1851.

[16] Giulio Casati, Felix Izrailev, and Luca Molinari. Scaling properties of the eigenvalue spacing distribution for band random matrices. *J. Phys. A*, 24(20):4755–4762, 1991. doi: 10.1088/0305-4470/24/20/011.

[17] Nixia Chen and Charles K. Smart. Random band matrix localization by scalar fluctuations, 2022. Preprint.

[18] Paul R. Chernoff. Essential self-adjointness of powers of generators of hyperbolic equations. *Journal of Functional Analysis*, 12:401–414, 1973.

[19] Giorgio Cipolloni, Ron Peled, Jeffrey Schenker, and Jacob Shapiro. Dynamical localization for random band matrices up to $W \ll N^{1/4}$. *Comm. Math. Phys.*, 405(3): Paper No. 82, 2024. doi: 10.1007/s00220-024-04948-1.

[20] Alain Comtet, Cécile Monthus, and Marc Yor. Exponential functionals of Brownian motion and disordered systems. *J. Appl. Probab.*, 35:255–271, 1998. doi: 10.1239/jap/1032192845.

[21] F. Constantinescu. The supersymmetric transfer matrix for linear chains with nondiagonal disorder. *J. Stat. Phys.*, 50(5-6):1167–1177, 1988. doi: 10.1007/BF01019159.

[22] Daryl J. Daley and David Vere-Jones. *An Introduction to the Theory of Point Processes. Vol. I: Elementary Theory and Methods*. Springer, New York, second edition, 2003.

[23] Daryl J. Daley and David Vere-Jones. *An Introduction to the Theory of Point Processes. Vol. II: General Theory and Structure*. Springer, New York, second edition, 2008.

[24] E. B. Davies. *Heat Kernels and Spectral Theory*. Cambridge University Press, Cambridge, 1989.

[25] Margherita Disertori. Density of states for GUE through supersymmetric approach. *Rev. Math. Phys.*, 16(9):1191–1225, 2004. doi: 10.1142/S0129055X04002229.

[26] Margherita Disertori and Mareike Lager. Density of states for random band matrices in two dimensions. *Ann. Henri Poincaré*, 18(7):2367–2413, 2017. doi: 10.1007/s00023-017-0572-3.

[27] Margherita Disertori, H. Pinson, and Thomas Spencer. Density of states for random band matrices. *Comm. Math. Phys.*, 232(1):83–124, 2002. doi: 10.1007/s00220-002-0733-0.

[28] Margherita Disertori, Thomas Spencer, and Martin R. Zirnbauer. Quasi-diffusion in a 3D supersymmetric hyperbolic sigma model. *Comm. Math. Phys.*, 300(2):435–486, 2010. doi: 10.1007/s00220-010-1117-5.

[29] Margherita Disertori, Martin Lohmann, and Sasha Sodin. The density of states of 1D random band matrices via a supersymmetric transfer operator. *J. Spectr. Theory*, 11(1):125–191, 2021.

[30] DLMF. NIST digital library of mathematical functions. `https://dlmf.nist.gov`, 2026. Accessed 2026-11-04.

[31] Reuben Drogin. Localization of one-dimensional random band matrices, 2025. Preprint.

[32] Sofiia Dubova and Kevin Yang. Quantum diffusion and delocalization in one-dimensional band matrices via the flow method. *J. Math. Phys.*, 67(3):033301, 2026. doi: 10.1063/5.0253873.

[33] Daniel Dufresne. The distribution of a perpetuity, with applications to risk theory and pension funding. *Scand. Actuar. J.*, 1990(1):39–79, 1990. doi: 10.1080/03461238.1990.10413872.

[34] Laure Dumaz and Cyril Labbé. The delocalized phase of the Anderson Hamiltonian in 1-d. *Ann. Probab.*, 51(3):805–839, 2023. doi: 10.1214/22-AOP1591.

[35] Laure Dumaz and Cyril Labbé. Localization crossover for the continuous Anderson Hamiltonian in 1-d. *Invent. Math.*, 235(2):345–440, 2024. doi: 10.1007/s00222-023-01225-1.

[36] K. B. Efetov. Supersymmetry method in localization theory. *Sov. Phys. JETP*, 55(3):514, 1982. Russian original: Zh. Eksp. Teor. Fiz. **82** (1982), 872.

[37] K. B. Efetov. Supersymmetry and theory of disordered metals. *Adv. Phys.*, 32(1):53–127, 1983. doi: 10.1080/00018738300101531.

[38] K. B. Efetov and A. I. Larkin. Kinetics of a quantum particle in a long metallic wire. *Sov. Phys. JETP*, 58(2):444–451, 1983. Russian original: Zh. Eksp. Teor. Fiz. **85** (1983), 764–778.

[39] Konstantin B. Efetov. *Supersymmetry in Disorder and Chaos*. Cambridge University Press, 1997.

[40] László Erdős and Antti Knowles. Quantum diffusion and eigenfunction delocalization in a random band matrix model. *Comm. Math. Phys.*, 303(2):509–554, 2011. doi: 10.1007/s00220-011-1204-2.

[41] László Erdős and Antti Knowles. Quantum diffusion and delocalization for band matrices with general distribution. *Ann. Henri Poincaré*, 12(7):1227–1319, 2011. doi: 10.1007/s00023-011-0104-5.

[42] László Erdős and Antti Knowles. The Altshuler–Shklovskii formulas for random band matrices I: the unimodular case. *Comm. Math. Phys.*, 333(3):1365–1416, 2015. doi: 10.1007/s00220-014-2119-5.

[43] László Erdős and Antti Knowles. The Altshuler–Shklovskii formulas for random band matrices II: the general case. *Ann. Henri Poincaré*, 16(3):709–799, 2015. doi: 10.1007/s00023-014-0333-5.

[44] László Erdős and Volodymyr Riabov. The zigzag strategy for random band matrices, 2025. Preprint.

[45] László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin. Delocalization and diffusion profile for random band matrices. *Comm. Math. Phys.*, 323(1):367–416, 2013. doi: 10.1007/s00220-013-1773-3.

[46] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986.

[47] Jiaqi Fan and Guangyi Zou. Poisson statistics of one-dimensional random band matrices, 2026. Preprint.

[48] Jiaqi Fan, Fan Yang, and Jun Yin. A block reduction method for random band matrices with general variance profiles, 2025. Preprint.

[49] Jiaqi Fan, Fan Yang, and Jun Yin. Localization lengths of power-law random band matrices, 2026. Preprint.

[50] William Feller. *An Introduction to Probability Theory and Its Applications. Vol. II*. John Wiley & Sons, second edition, 1971.

[51] Yan V. Fyodorov. Negative moments of characteristic polynomials of random matrices: Ingham–Siegel integral as an alternative to Hubbard–Stratonovich transformation. *Nuclear Phys. B*, 621(3):643–674, 2002. doi: 10.1016/S0550-3213(01)00508-9.

[52] Yan V. Fyodorov. On Hubbard–Stratonovich transformations over hyperbolic domains. *J. Phys.: Condens. Matter*, 17(20):S1915–S1928, 2005. doi: 10.1088/0953-8984/17/20/018.

[53] Yan V. Fyodorov and Alexander D. Mirlin. Scaling properties of localization in random band matrices: a σ-model approach. *Phys. Rev. Lett.*, 67(18):2405–2409, 1991. doi: 10.1103/PhysRevLett.67.2405.

[54] Yan V. Fyodorov and Alexander D. Mirlin. Analytical derivation of the scaling law for the inverse participation ratio in quasi-one-dimensional disordered systems. *Phys. Rev. Lett.*, 69(7):1093–1096, 1992. doi: 10.1103/PhysRevLett.69.1093. Erratum: Phys. Rev. Lett. **69** (1992), 1832.

[55] Yan V. Fyodorov and Alexander D. Mirlin. Level-to-level fluctuations of the inverse participation ratio in finite quasi 1D disordered systems. *Phys. Rev. Lett.*, 71(3): 412–415, 1993. doi: 10.1103/PhysRevLett.71.412.

[56] Yan V. Fyodorov and Alexander D. Mirlin. Statistical properties of eigenfunctions of random quasi 1D one-particle Hamiltonians. *Int. J. Mod. Phys. B*, 8(27):3795–3842, 1994. doi: 10.1142/S0217979294001640.

[57] Yan V. Fyodorov and Alexander D. Mirlin. Mesoscopic fluctuations of eigenfunctions and level-velocity distribution in disordered metals. *Phys. Rev. B*, 51(19): 13403–13409, 1995. doi: 10.1103/PhysRevB.51.13403.

[58] Yan V. Fyodorov and Alexander D. Mirlin. Statistical properties of random banded matrices with strongly fluctuating diagonal elements. *Phys. Rev. B*, 52(16):R11580–R11583, 1995. doi: 10.1103/PhysRevB.52.R11580.

[59] Yan V. Fyodorov and Eugene Strahov. An exact formula for general spectral correlation function of random Hermitian matrices. *J. Phys. A*, 36(12):3203–3213, 2003. doi: 10.1088/0305-4470/36/12/320.

[60] Yan V. Fyodorov, Y. Wei, and Martin R. Zirnbauer. Hyperbolic Hubbard–Stratonovich transformation made rigorous. *J. Math. Phys.*, 49(5):053507, 2008. doi: 10.1063/1.2917066.

[61] Subhroshekhar Ghosh and Yuval Peres. Rigidity and tolerance in point processes: Gaussian zeros and Ginibre eigenvalues. *Duke Math. J.*, 166(10):1789–1858, 2017. doi: 10.1215/00127094-2017-0002.

[62] David Gilbarg and Neil S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Springer, Berlin, 2001.

[63] Gian Michele Graf and Alessio Vaghi. A remark on the estimate of a determinant by Minami. *Lett. Math. Phys.*, 79(1):17–22, 2007. arXiv:math-ph/0604033.

[64] Alexander Grigor'yan. *Heat Kernel and Analysis on Manifolds*, volume 47 of *AMS/IP Studies in Advanced Mathematics*. American Mathematical Society, Providence, RI, 2009.

[65] Batu Güneysu. *Covariant Schrödinger Semigroups on Riemannian Manifolds*, volume 264 of *Operator Theory: Advances and Applications*. Birkhäuser, 2017.

[66] Sigurdur Helgason. *Groups and Geometric Analysis: Integral Geometry, Invariant Differential Operators, and Spherical Functions*. Academic Press, Orlando, FL, 1984.

[67] Peter D. Hislop and M. Krishna. On the local eigenvalue statistics for random band matrices in the localization regime. *J. Stat. Phys.*, 187(3):Paper No. 26, 2022. doi: 10.1007/s10955-022-02923-5.

[68] Loo-Keng Hua. Inequalities involving determinants. *Acta Math. Sinica*, 5:463–470, 1955. English translation in Amer. Math. Soc. Transl. Ser. 2 **32** (1963), 265–272.

[69] W. Hunziker. Distortion analyticity and molecular resonance curves. *Ann. Inst. H. Poincaré Phys. Théor.*, 45:339–358, 1986.

[70] Kiyosi Itô and Henry P. McKean, Jr. *Diffusion Processes and their Sample Paths*, volume 125 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 1965.

[71] D. A. Ivanov, P. M. Ostrovsky, and M. A. Skvortsov. Correlations of the local density of states in quasi-one-dimensional wires. *Phys. Rev. B*, 79(20):205108, 2009. doi: 10.1103/PhysRevB.79.205108.

[72] Tosio Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics. Springer, Berlin, 1995. doi: 10.1007/978-3-642-66282-9. Reprint of the 1980 edition.

[73] Yitzhak Katznelson. *An Introduction to Harmonic Analysis*. Cambridge University Press, Cambridge, third edition, 2004.

[74] S. Kettemann. Symmetry dependence of localization in quasi-one-dimensional disordered wires. *Phys. Rev. B*, 62(20):R13282–R13285, 2000. doi: 10.1103/PhysRevB.62.R13282.

[75] Stefan Kettemann. Exploring level statistics from quantum chaos to localization with the autocorrelation function of spectral determinants. *Phys. Rev. B*, 59(7): 4799–4802, 1999. doi: 10.1103/PhysRevB.59.4799.

[76] Mario Kieburg. On the Efetov–Wegner terms by diagonalizing a Hermitian supermatrix. *J. Phys. A*, 44(28):285210, 2011. doi: 10.1088/1751-8113/44/28/285210.

[77] Rowan Killip and Mihai Stoiciu. Eigenvalue statistics for CMV matrices: from Poisson to clock via random matrix ensembles. *Duke Math. J.*, 146(3):361–399, 2009. doi: 10.1215/00127094-2009-001.

[78] Abel Klein, Fabio Martinelli, and J. Fernando Perez. A rigorous replica trick approach to Anderson localization in one dimension. *Comm. Math. Phys.*, 106(4): 623–633, 1986.

[79] I. V. Kolokolov. The method of functional integration for one-dimensional localization, higher correlators, and the average current flowing in a mesoscopic ring in an arbitrary magnetic field. *Sov. Phys. JETP*, 76(6):1099, 1993. Russian original: Zh. Eksp. Teor. Fiz. **103** (1993), 2196.

[80] V. E. Kravtsov and A. D. Mirlin. Level statistics in a metallic sample: corrections to the Wigner–Dyson distribution. *JETP Lett.*, 60:656, 1994. Russian original: Pis'ma Zh. Eksp. Teor. Fiz. **60** (1994), 645.

[81] Evgenij Kritchevski, Benedek Valkó, and Bálint Virág. The scaling limit of the critical one-dimensional random Schrödinger operator. *Comm. Math. Phys.*, 314(3): 775–806, 2012. doi: 10.1007/s00220-012-1537-5.

[82] Thomas G. Kurtz. Extensions of Trotter's operator semigroup approximation theorems. *J. Funct. Anal.*, 3:354–375, 1969.

[83] Thomas G. Kurtz. A limit theorem for perturbed operator semigroups with applications to random evolutions. *J. Funct. Anal.*, 12:55–67, 1973.

[84] Dimitry A. Leites. Introduction to the theory of supermanifolds. *Russian Math. Surveys*, 35(1):1–64, 1980. doi: 10.1070/RM1980v035n01ABEH001545.

[85] J.-L. Lions. *Équations Différentielles Opérationnelles et Problèmes aux Limites*. Springer, Berlin, 1961.

[86] Dang-Zheng Liu and Guangyi Zou. Edge statistics for random band matrices, 2023. Preprint.

[87] M. Marinho and T. Micklitz. Spectral correlations in Anderson insulating wires. *Phys. Rev. B*, 97:041406(R), 2018. arXiv:1711.00151.

[88] Hiroyuki Matsumoto and Marc Yor. Exponential functionals of Brownian motion, I: Probability laws at fixed time. *Probab. Surv.*, 2:312–347, 2005. doi: 10.1214/154957805100000159.

[89] F. Gustav Mehler. Ueber die Entwicklung einer Function von beliebig vielen Variablen nach Laplaceschen Functionen höherer Ordnung. *J. Reine Angew. Math.*, 66:161–176, 1866.

[90] T. Micklitz. Spectral correlations in finite-size Anderson insulators. *Phys. Rev. B*, 93: 094201, 2016. arXiv:1611.07271.

[91] T. Micklitz, C. A. Müller, and A. Altland. Strong Anderson localization in cold atom quantum quenches. *Phys. Rev. Lett.*, 112:110602, 2014. arXiv:1311.2268.

[92] Nariyuki Minami. Local fluctuation of the spectrum of a multidimensional Anderson tight binding model. *Comm. Math. Phys.*, 177(3):709–725, 1996.

[93] A. D. Mirlin and Y. V. Fyodorov. The statistics of eigenvector components of random band matrices: analytical results. *J. Phys. A*, 26(12):L551–L558, 1993. doi: 10.1088/0305-4470/26/12/012.

[94] Alexander D. Mirlin. Statistics of energy levels and eigenfunctions in disordered systems. *Phys. Rep.*, 326:259–382, 2000. doi: 10.1016/S0370-1573(99)00091-5.

[95] Alexander D. Mirlin. Statistics of energy levels and eigenfunctions in disordered and chaotic systems: supersymmetry approach. In G. Casati, I. Guarneri, and U. Smilansky, editors, *New Directions in Quantum Chaos (Proceedings of the International School of Physics "Enrico Fermi", Course CXLIII)*, pages 223–298. IOS Press, Amsterdam, 2000.

[96] J. Müller-Hill and Martin R. Zirnbauer. Equivalence of domains for hyperbolic Hubbard–Stratonovich transformations. *J. Math. Phys.*, 52(5):053506, 2011. doi: 10.1063/1.3585672.

[97] Fumihiko Nakano. Shape of eigenvectors for the decaying potential model. *Ann. Henri Poincaré*, 24(3):871–893, 2023. doi: 10.1007/s00023-022-01222-9.

[98] L. Nirenberg. An abstract form of the nonlinear Cauchy–Kowalewski theorem. *J. Differential Geom.*, 6:561–576, 1972.

[99] Takaaki Nishida. A note on a theorem of Nirenberg. *J. Differential Geom.*, 12:629–633, 1977.

[100] Sheehan Olver and Andrew Swan. Evidence of the Poisson/Gaudin–Mehta phase transition for band matrices on global scales. *Random Matrices Theory Appl.*, 7(2): 1850002, 2018. doi: 10.1142/S2010326318500028.

[101] L. V. Ovsjannikov. A nonlinear Cauchy problem in a scale of Banach spaces. *Dokl. Akad. Nauk SSSR*, 200:789–792, 1971.

[102] Chris Paige, George P. H. Styan, Bo-Ying Wang, and Fuzhen Zhang. Hua's matrix equality and Schur complements. *Int. J. Inf. Syst. Sci.*, 4(1):124–135, 2008.

[103] Giorgio Parisi and Nicolas Sourlas. Random magnetic fields, supersymmetry, and negative dimensions. *Phys. Rev. Lett.*, 43:744–745, 1979.

[104] Ron Peled, Jeffrey Schenker, Mira Shamis, and Sasha Sodin. On the Wegner orbital model. *Int. Math. Res. Not. IMRN*, 2019(4):1030–1058, 2019.

[105] Ross G. Pinsky. *Positive Harmonic Functions and Diffusion*, volume 45 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.

[106] Thomas Ransford. *Potential Theory in the Complex Plane*. Cambridge University Press, Cambridge, 1995.

[107] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-Adjointness*. Academic Press, New York, 1975.

[108] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, third edition, 1999. doi: 10.1007/978-3-662-06400-9.

[109] Ben Rifkind and Bálint Virág. Eigenvectors of the 1-dimensional critical random Schrödinger operator. *Geom. Funct. Anal.*, 28(5):1394–1419, 2018. arXiv:1605.00118.

[110] Mitchell J. Rothstein. Integration on noncompact supermanifolds. *Trans. Amer. Math. Soc.*, 299(1):387–396, 1987. doi: 10.1090/S0002-9947-1987-0869418-5.

[111] Lothar Schäfer and Franz Wegner. Disordered system with n orbitals per site: Lagrange formulation, hyperbolic symmetry, and Goldstone modes. *Z. Phys. B*, 38 (2):113–126, 1980. doi: 10.1007/BF01598751.

[112] Jeffrey Schenker. Eigenvector localization for random band matrices with power law band width. *Comm. Math. Phys.*, 290(3):1065–1097, 2009. doi: 10.1007/s00220-009-0798-0.

[113] Albert Schwarz and Oleg Zaboronsky. Supersymmetry and localization. *Comm. Math. Phys.*, 183:463–476, 1997.

[114] T. H. Seligman, J. J. M. Verbaarschot, and M. R. Zirnbauer. Spectral fluctuation properties of Hamiltonian systems: the transition region between order and chaos. *J. Phys. A*, 18(14):2751–2770, 1985. doi: 10.1088/0305-4470/18/14/026.

[115] Mariya Shcherbina and Tatyana Shcherbina. Transfer matrix approach to 1d random band matrices: density of states. *J. Stat. Phys.*, 164:1233–1260, 2016. doi: 10.1007/s10955-016-1593-x.

[116] Mariya Shcherbina and Tatyana Shcherbina. Characteristic polynomials for 1D random band matrices from the localization side. *Comm. Math. Phys.*, 351:1009–1044, 2017. doi: 10.1007/s00220-017-2849-2.

[117] Mariya Shcherbina and Tatyana Shcherbina. Universality for 1d random band matrices: sigma-model approximation. *J. Stat. Phys.*, 172:627–664, 2018. doi: 10.1007/s10955-018-1969-1.

[118] Mariya Shcherbina and Tatyana Shcherbina. Transfer operator approach to 1d random band matrices. In *Proceedings of the International Congress of Mathematicians (Rio de Janeiro 2018)*, pages 2687–2707. World Scientific, 2019. doi: 10.1142/9789813272880_0156. arXiv:1905.08252.

[119] Mariya Shcherbina and Tatyana Shcherbina. Universality for 1d random band matrices. *Comm. Math. Phys.*, 385(2):667–716, 2021. doi: 10.1007/s00220-021-04135-6.

[120] Mariya Shcherbina and Tatyana Shcherbina. Characteristic polynomials of non-Hermitian random band matrices near the threshold, 2026. Preprint.

[121] Tatyana Shcherbina. On the second mixed moment of the characteristic polynomials of 1D band matrices. *Comm. Math. Phys.*, 328(1):45–82, 2014. doi: 10.1007/s00220-014-1947-7.

[122] Tatyana Shcherbina. Characteristic polynomials for random band matrices near the threshold. *J. Stat. Phys.*, 179(4):920–944, 2020. doi: 10.1007/s10955-020-02567-3.

[123] Tatyana Shcherbina. SUSY transfer matrix approach for the real symmetric 1d random band matrices. *Electron. J. Probab.*, 27:1–29, 2022. doi: 10.1214/22-EJP747.

[124] M. A. Shubin. Spectral theory of elliptic operators on non-compact manifolds. *Astérisque*, 207:35–108, 1992.

[125] Barry Simon. *Functional Integration and Quantum Physics*. AMS Chelsea Publishing, Providence, RI, second edition, 2005.

[126] M. A. Skvortsov and P. M. Ostrovsky. Local correlations of different eigenfunctions in a disordered wire. *JETP Lett.*, 85(1):72–76, 2007. doi: 10.1134/S0021364007010158. Russian original: Pis'ma Zh. Eksp. Teor. Fiz. **85** (2007), 79.

[127] Sasha Sodin. The spectral edge of some random band matrices. *Ann. of Math. (2)*, 172(3):2223–2251, 2010. doi: 10.4007/annals.2010.172.2223.

[128] Thomas Spencer. SUSY statistical mechanics and random band matrices. In *Quantum Theory from Small to Large Scales*, volume 95 of *Lecture Notes of the Les Houches Summer School*, pages 467–518. Oxford University Press, 2012. doi: 10.1093/acprof:oso/9780199652495.003.0007.

[129] Thomas Spencer. Random banded and sparse matrices. In *The Oxford Handbook of Random Matrix Theory*, pages 470–488. Oxford University Press, 2015. doi: 10.1093/oxfordhb/9780198744191.013.23.

[130] Béla Sz.-Nagy, Ciprian Foiaş, Hari Bercovici, and László Kérchy. *Harmonic Analysis of Operators on Hilbert Space*. Universitext. Springer, New York, second edition, 2010.

[131] H. F. Trotter. Approximation of semi-groups of operators. *Pacific J. Math.*, 8:887–919, 1958.

[132] V. S. Varadarajan. *Supersymmetry for Mathematicians: An Introduction*, volume 11 of *Courant Lecture Notes in Mathematics*. American Mathematical Society, 2004.

[133] J. J. M. Verbaarschot, H. A. Weidenmüller, and M. R. Zirnbauer. Grassmann integration in stochastic quantum physics: the case of compound-nucleus scattering. *Phys. Rep.*, 129(6):367–438, 1985. doi: 10.1016/0370-1573(85)90070-5.

[134] Franz Wegner. The mobility edge problem: continuous symmetry and a conjecture. *Z. Phys. B*, 35(3):207–210, 1979. doi: 10.1007/BF01319839.

[135] Franz Wegner. *Supermathematics and its Applications in Statistical Physics: Grassmann Variables and the Method of Supersymmetry*. Lecture Notes in Physics. Springer, Berlin, Heidelberg, 2016. doi: 10.1007/978-3-662-49170-6.

[136] Hassler Whitney. Differentiable even functions. *Duke Math. J.*, 10:159–160, 1943.

[137] Fan Yang and Jun Yin. Random band matrices in the delocalized phase, III: averaging fluctuations. *Probab. Theory Related Fields*, 179(1-2):451–540, 2021. doi: 10.1007/s00440-020-01013-5.

[138] Horng-Tzer Yau and Jun Yin. Delocalization of one-dimensional random band matrices, 2025. Preprint.

[139] Martin R. Zirnbauer. Riemannian symmetric superspaces and their origin in random-matrix theory. *J. Math. Phys.*, 37(10):4986–5018, 1996. doi: 10.1063/1.531675.

[140] Martin R. Zirnbauer. The supersymmetry method of random matrix theory. In *Encyclopedia of Mathematical Physics*, volume 5, pages 151–160. Elsevier, Amsterdam, 2006.