Random partitions and instanton counting

Andrei Okounkov*

Abstract. We summarize the connection between random partitions and $\mathcal{N} = 2$ supersymmetric gauge theories in 4 dimensions and indicate how this relation extends to higher dimensions.

Mathematics Subject Classification (2000). Primary 81T13; Secondary 14J60.

1. Introduction

1.1. Random partitions. A partition of n is a monotone sequence

$$\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq 0)$$

of nonnegative integers with sum n. The number n is denoted $|\lambda|$ and called the size of λ. A geometric object associated to a partition is its *diagram*; it contains λ_1 squares in the first row, λ_2 squares in the second row and so on. An example, flipped and rotated by 135° can be seen in Figure 1. Partitions naturally label many basic objects

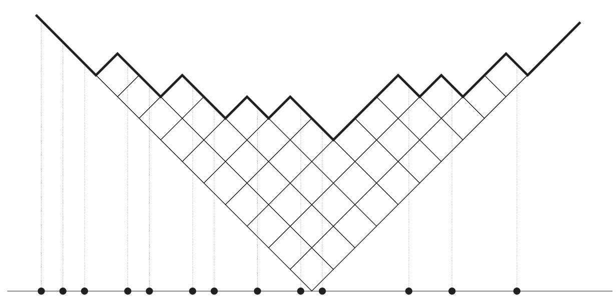


Figure 1. The diagram of $\lambda = (10, 8, 7, 4, 4, 3, 2, 2, 1, 1)$, flipped and rotated by 135°. Bullets indicate the points of $\mathfrak{S}(\lambda)$. The profile of λ is plotted in bold.

in mathematics and physics, such as e.g. conjugacy classes and representations of the symmetric group $S(n)$, and very often appear as modest summation ranges or indices. A simple but fruitful change of perspective, which I wish to stress here, is to treat sums over partitions probabilistically, that is, treat them as expectations of some functions of a random partition.

*The author thanks Packard Foundation for partial financial support.

Proceedings of the International Congress
of Mathematicians, Madrid, Spain, 2006

A survey of the theory of random partitions may be found in [33]. Of the several natural measures on partitions, the Plancherel measure

$$\mathfrak{M}_{\text{Planch}}(\lambda) = \frac{(\dim \lambda)^2}{n!}, \quad |\lambda| = n, \tag{1}$$

stands out as the one with deepest properties and widest applications. Here $\dim \lambda$ is the dimension of the corresponding representation of $S(n)$. This is a probability measure on partitions of n. It can be viewed as a distinguished discretization of the GUE ensemble of the random matrix theory. Namely, a measure on partitions can be made a point process on a lattice by placing particles in positions

$$\mathfrak{S}(\lambda) = \left\{\lambda_i - i + \tfrac{1}{2}\right\} \subset \mathbb{Z} + \tfrac{1}{2}.$$

Figure 1 illustrates the geometric meaning of this transformation. An important theme of recent research was to understand how and why for a Plancherel random partition of $n \to \infty$ the particles $\mathfrak{S}(\lambda)$ behave like the eigenvalues of a random Hermitian matrix. See [3], [4], [14], [31] and e.g. [15], [32], [33] for a survey.

In these notes, we consider a different problem, namely, the behavior of (1) in a *periodic potential*, that is, additionally weighted by a multiplicative periodic function of the particles' positions. This leads to new phenomena and new applications. As we will see, the partition function of $\mathfrak{M}_{\text{Planch}}$ in a periodic potential is closely related to Nekrasov partition function from supersymmetric gauge theory. This relationship will be reviewed in detail in Section 2 and its consequences may be summarized as follows.

1.2. Instanton counting. In 1994, Seiberg and Witten proposed an exact description of the low-energy behavior of certain supersymmetric gauge theories [38], [39]. In spite of the enormous body of research that this insight has generated, only a modest progress was made towards its gauge-theoretic derivation. This changed in 2002, when Nekrasov proposed in [28] a physically meaningful and mathematically rigorous definition of the regularized partition function Z for supersymmetric gauge theories in question.

Supersymmetry makes the gauge theory partition function the partition function of a *gas of instantons*. Nekrasov's idea was to use *equivariant integration* with respect to the natural symmetry group in lieu of a long-distance cut-off for the instanton gas. He conjectured that as the regularization parameter $\varepsilon \to 0$

$$\ln Z \sim -\frac{1}{\varepsilon^2} \mathcal{F}$$

where the free energy $\mathcal{F}$ expressed by the Seiberg–Witten formula in terms of periods of a certain differential dS on a certain algebraic curve C.

This conjecture was proven in 2003 by Nekrasov and the author for a list of gauge theories with gauge group $U(r)$, namely, pure gauge theory, theories with matter fields in fundamental and adjoint representations of the gauge group, as well as 5-dimensional theory compactified on a circle [29]. Simultaneously, independently, and using completely different ideas, the formal power series version of Nekrasov's conjecture was proven for the pure $U(r)$-theory by Nakajima and Yoshioka [25]. The methods of [29] were applied to classical gauge groups in [30] and to the 6-dimensional gauge theory compactified on a torus in [11]. Another algebraic approach, which works for pure gauge theory with any gauge group, was developed by Braverman [5] and Braverman and Etingof [6].

In these notes, we outline the results of [29] in the simplest, yet fundamental, case of pure gauge theory. As should be obvious from the title, the main idea is to treat the gauge theory partition function Z as the partition function of an ensemble of random partitions. The $\varepsilon \to 0$ limit turns out to be the *thermodynamic limit* in this ensemble. What emerges in this limit is a nonrandom *limit shape*, an example of which may be seen in Figure 6. This is a form of the law of large numbers, analogous, for example, to Wigner's semicircle law for the spectrum of a large random matrix. The limit shape is characterized as the unique minimizer $\psi_\star$ of a certain convex functional $\mathcal{S}(\psi)$, leading to

$$\mathcal{F} = \min \mathcal{S}.$$

We solve the variational problem explicitly and the limit shape turns out to be an algebraic curve C is disguise. Namely, the limit shape is essentially the graph of the function

$$\Re \int_{x_0}^{x} dS,$$

where dS is the Seiberg–Witten differential. Thus all ingredients of the answer appear very naturally in the proof.

Random matrix theory and philosophy had many successes in mathematics and physics. Here we have an example when random partitions, while structurally resembling random matrices, offer several advantages. First, the transformation into a random partition problem is geometrically natural and exact. Second, the discretization inherent in partitions regularizes several analytic issues. For further examples along these lines the reader may consult [33].

1.3. Higher dimensions. The translation of the gauge theory problem into a random partition problem is explained in Section 2. In Section 3, we analyze the latter problem, in particular, derive and solve the variational problem for the limit shape. Section 4 summarizes parallel results for 3-dimensional partitions, where similar algebraic properties of limit shapes are now proven in great generality.

The surprising fact that free energy $\mathcal{F}$ is given in terms of periods of a hidden algebraic curve C is an example of *mirror symmetry*. A general program of interpreting mirror partners as limit shapes was initiated in [34]. Known results about the limit shapes of periodically weighted 3-dimensional partitions, together with the conjectural equality of Gromov–Witten and Donaldson–Thomas theories of projective algebraic 3-folds [22] can be interpreted as a verification of this program for toric Calabi–Yau 3-folds. See [35] for an introduction to these ideas.

Note that something completely different is expected to happen in dimensions > 3, where the behavior of both random interfaces and Gromov–Witten invariants changes qualitatively.

2. The gauge theory problem

2.1. Instantons. We begin by recalling some basic facts, see [9] for an excellent mathematical treatment and [8], [10], [44] for a physical one. This will serve as motivation for the introduction of Nekrasov's partitions function in (8) below.

In gauge theories, interactions are transmitted by gauge fields, that is, unitary connections on appropriate vector bundles. In coordinates, these are matrix-valued functions $A_i(x)$ that define covariant derivatives

$$\nabla_i = \frac{\partial}{\partial x_i} + A_i(x), \quad A_i^* = -A_i .$$

We consider the most basic case of the trivial bundle $\mathbb{R}^4 \times \mathbb{C}^r$ over the flat Euclidean space-time $\mathbb{R}^4$, where such coordinate description is global.

The natural (Yang–Mills) energy functional for gauge fields is L^2-norm squared $\|F\|^2$ of the curvature

$$F = \sum \left[\nabla_i, \nabla_j\right] dx_i \wedge dx_j .$$

The path integral in quantum gauge theory then takes the form

$$\int_{\text{connections}/\mathcal{G}} \mathcal{D}A \;\; \exp\left(-\beta \, \|F\|^2\right) \times \cdots , \tag{2}$$

where dots stand for terms involving other fields of the theory and $\mathcal{G}$ is the group of gauge transformations $g : \mathbb{R}^4 \to U(r)$ acting by

$$\nabla \mapsto g \, \nabla \, g^{-1} .$$

In these notes, we will restrict ourselves to pure gauge theory, which is already quite challenging due to the complicated form of the energy. A parallel treatment of certain matter fields can be found in [29].

Our goal is to study (2) as function of the parameter β (and boundary conditions at infinity, see below). A head-on probabilistic approach to this problem would be to make it a theory of many interacting random matrices through a discretization of space-time. This is a fascinating topic about which I have nothing to say. In a different direction, when $\beta \gg 0$, the minima of $\|F\|^2$ should dominate the integral.

In *supersymmetric* gauge theory, there is a way to make such approximation exact, thereby reducing the path integral to the following finite-dimensional integrals.

Local minima of $\|F\|^2$ are classified by a topological invariant $c_2 \in \mathbb{Z}$,

$$c_2 = \frac{1}{8\pi^2} \int_{\mathbb{R}^4} \operatorname{tr} F^2,$$

called *charge*, and satisfy a system of first order PDEs

$$F \pm \star F = 0, \tag{3}$$

where $\star$ is the Hodge star operator on 2-forms on $\mathbb{R}^4$. With the plus sign, (3) corresponds to $c_2 > 0$ and is called the anti-self-duality equation. Its solutions are called *instantons*. Minima with $c_2 < 0$ are obtained by reversing the orientation of $\mathbb{R}^4$.

The ASD equations (3) are conformally invariant and can be transported to a punctured 4-sphere $S^4 = \mathbb{R}^4 \cup \{\infty\}$ via stereographic projection. From the removable singularities theorem of Uhlenbeck it follows that any instanton on $\mathbb{R}^4$ extends, after a gauge transformation, to an instanton on S^4. Thus we can talk about the value of an instanton at infinity.

Let $\mathcal{G}_0$ be the group of maps $g : S^4 \to U(r)$ such that $g(\infty) = 1$. Modulo $\mathcal{G}_0$, instantons on S^4 with $c_2 = n$ are parametrized by a smooth manifold $\mathcal{M}(r, n)$ of real dimension $4rn$. Naively, one would like the contribution from charge n instantons to (2) to be the volume of $\mathcal{M}(r, n)$ in a natural symplectic structure. However, $\mathcal{M}(r, n)$ is noncompact (and its volume is infinite) for two following reasons.

Approximately, an element of $\mathcal{M}(r, n)$ can be imagined as a nonlinear superposition of n instantons of charge 1. Some of those may become point-like, i.e. their curvature may concentrate in a δ-function spike, while others may wander off to infinity. A partial compactification of $\mathcal{M}(r, n)$, constructed by Uhlenbeck, which replaces point-like instanton by just points of $\mathbb{R}^4$, takes care of the first problem but not the second. Nekrasov's idea was to use *equivariant integration* to regularize the instanton contributions.

2.2. Equivariant regularization. The group

$$K = \mathrm{SU}(2) \times \mathrm{SU}(r)$$

acts on $\mathcal{M}(r, n)$ by rotations of $\mathbb{R}^4 = \mathbb{C}^2$ and constant gauge transformation, respectively. Our plan is to use this action for regularization. Let us start with the following simplest example: suppose we want to regularize the volume of $\mathbb{R}^2$. A gentle way to do it is to introduce a Gaussian well

$$\int_{\mathbb{R}^2} e^{-t\pi(x^2+y^2)} dx\, dy = \frac{1}{t}, \quad \Re t \geq 0 \tag{4}$$

and thus an effective cut-off at the $|t|^{-1/2}$ scale. Note that the Hamiltonian flow on $\mathbb{R}^2$ generated by $H = \frac{1}{2}(x^2 + y^2)$ with respect to the standard symplectic form

$\omega = dx \wedge dy$ is rotation about the origin with angular velocity one. This makes (4) a simplest instance of the Atiyah–Bott–Duistermaat–Heckman *equivariant localization* formula [2]. We will use localization in the following complex form.

Let $T = \mathbb{C}^*$ act on a complex manifold X with isolated fixed points X^T. Suppose that the action of $U(1) \subset T$ is generated by a Hamiltonian H with respect to a symplectic form ω. Then

$$\int_X e^{\omega - 2\pi t H} = \sum_{x \in X^T} \frac{e^{-2\pi t H(x)}}{\det t|_{T_x X}}, \tag{5}$$

where t should be viewed as an element of $\mathrm{Lie}(T) \cong \mathbb{C}$, an so it acts in the complex tangent space $T_x X$ to a fixed point $x \in X$. While (5) is normally stated for compact manifolds X, example (4) shows that with care it can work for noncompact ones, too. Scaling both ω and H to zero, we get from (5) a formal expression

$$\int_X 1 \stackrel{\text{def}}{=} \sum_{x \in X^T} \frac{1}{\det t|_{T_x X}}, \tag{6}$$

which does not depends on the symplectic form and vanishes if X is compact.

A theorem of Donaldson identifies instantons with *holomorphic bundles* on $\mathbb{C}^2 = \mathbb{R}^4$ and thus gives a complex description of $\mathcal{M}(r, n)$. Concretely, $\mathcal{M}(r, n)$ is the moduli space of rank r holomorphic bundles $\mathcal{E} \to \mathbb{CP}^2$ with given 2nd Chern class $c_2(\mathcal{E}) = n$ and a given trivialization along the line

$$L_\infty = \mathbb{CP}^2 \setminus \mathbb{C}^2$$

at infinity. Note that existence of such trivialization implies that $c_1(\mathcal{E}) = 0$. A similar but larger moduli space $\overline{\mathcal{M}}(r, n)$ of *torsion-free sheaves*, see e.g. [12], [24], is a smooth partial compactification of $\mathcal{M}(r, n)$.

The complexification of K

$$K_{\mathbb{C}} = \mathrm{SL}(2) \times \mathrm{SL}(r)$$

acts on $\overline{\mathcal{M}}(r, n)$ by operating on $\mathbb{C}^2$ and changing the trivialization at infinity. Equivariant localization with respect to a general $t \in \mathrm{Lie}(K)$

$$t = (\mathrm{diag}(-i\varepsilon, i\varepsilon), \mathrm{diag}(ia_1, \dots, ia_r)) \tag{7}$$

combines the two following effects. First, it introduces a spatial cut-off parameter ε as in (4). Second, it introduces dependence on the instanton's behavior at infinity through the parameters a_i. While the first factor in K works to shepherd run-away instantons back to the origin, the second works to break the gauge invariance at infinity. In supersymmetric gauge theories, the parameters a_i correspond to the vacuum expectation of the Higgs field and thus are responsible for masses of gauge bosons. In short, they are live physical parameters.

2.3. Nekrasov partition function. We are now ready to introduce our main object of study, the partition function of the pure ($\mathcal{N} = 2$ supersymmetric) $U(r)$ gauge theory:

$$Z(\varepsilon; a_1, \ldots, a_r; \Lambda) = Z_{\text{pert}} \sum_{n \geq 0} \Lambda^{2rn} \int_{\overline{\mathcal{M}}(r,n)} 1, \tag{8}$$

where the integral is defined by (6) applied to (7),

$$\Lambda = \exp(-4\pi^2 \beta / r),$$

and Z_{pert} is a certain perturbative factor to be discussed below. The series in (8) is denoted Z_{inst}. Because of factorials in denominators, see (24), Z_{inst} converges whenever we avoid zero denominators, that is, on the complement of

$$a_i - a_j \equiv 0 \mod \varepsilon. \tag{9}$$

In essence, these factorials are there because the instantons are unordered. Also note that Z is an even function of ε and a symmetric function of the a_i's.

Since by our regularization rule

$$\operatorname{vol} \mathbb{R}^4 = \int_{\mathbb{R}^4} 1 = \frac{1}{\varepsilon^2},$$

we may expect that as $\varepsilon \to 0$

$$\ln Z(\varepsilon; a; \Lambda) \sim -\frac{1}{\varepsilon^2} \mathcal{F}(a; \Lambda),$$

where $\mathcal{F}$ is the *free energy*. At first, the poles (9) of Z_{inst}, which are getting denser and denser, may look like a problem. Indeed, poles of multiplicity $O(\varepsilon^{-1})$ may affect the free energy, but it is a question of competition with the other terms in Z_{inst}, which the pole-free terms win if $|a_i - a_j| \gg 0$. As a result, either by passing to a subsequence of ε, or by restricting summation in Z_{inst} to the relevant pole-free terms, we obtain a limit

$$\mathcal{F}_{\text{inst}} = -\lim \varepsilon^2 Z_{\text{inst}},$$

which is analytic and monotone far enough from the walls of the Weyl chambers. Recall that Weyl chambers for $\mathrm{SU}(r)$ are the $r!$ cones obtained from

$$\mathsf{C}_+ = \left\{ a_1 > a_2 > \cdots > a_r, \ \sum a_i = 0 \right\}$$

by permuting the coordinates. As $|a_i - a_j|$ get small, poles do complicate the asymptotics. This is the origin of cuts in the analytic function $\mathcal{F}(a)$, $a_i \in \mathbb{C}$.

Nekrasov conjectured in [28] that the free energy $\mathcal{F}$ is the *Seiberg–Witten prepotential*, first obtained in [38], [39] through entirely different considerations. It is defined in terms of a certain family of algebraic curves.

2.4. Seiberg–Witten geometry. In the affine space of complex polynomials of the form $P(z) = z^r + O(z^{r-2})$ consider the open set $\boldsymbol{U}$ of polynomials such that

$$P(z) = \pm 2\Lambda^r \tag{10}$$

has $2r$ distinct roots. Over $\boldsymbol{U}$, we have a g-dimensional family of complex algebraic curves C of genus $g = r - 1$ defined by

$$\Lambda^r \left(w + \frac{1}{w} \right) = P(z), \quad P \in \boldsymbol{U}. \tag{11}$$

The curve (11) is compactified by adding two points $\partial C = \{w = 0, \infty\}$.

Let $\mathsf{M} \subset \boldsymbol{U}$ be the set of $P(z)$ for which all roots of (10) are real. The corresponding curves C are called *maximal* and play a special role, see e.g. [40]. They arise, for example, as spectral curves of a periodic Toda chain [41]. A maximal curve C has r real ovals, as illustrated in Figure 2. Note that for $z \in \mathbb{R}$, w is either real or lies on the unit circle $|w| = 1$.

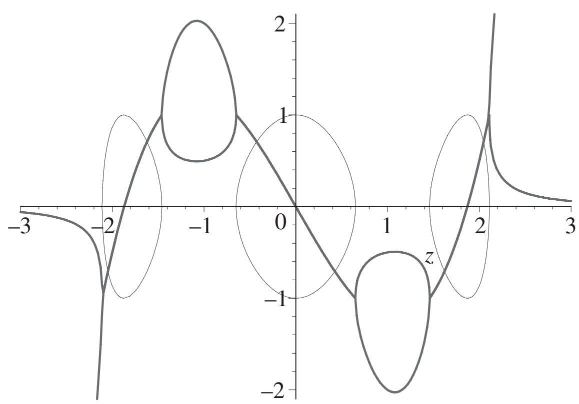


Figure 2. $\Re w$ (bold) and $\Im w$ for $w + 1/w = z^3 - 3.5\,z$ and $z \in \mathbb{R}$.

The intervals $P^{-1}([-2\Lambda^r, 2\Lambda^r]) \subset \mathbb{R}$ on which $|w| = 1$ are called *bands*. The intervals between the bands are called *gaps*. The smaller (in absolute value) root w of the equation (11) can be unambiguously defined for $z \in \mathbb{C} \setminus \{\text{bands}\}$. On the corresponding sheet of the Riemann surface of w, we define cycles

$$\alpha_i \in H_1(C - \partial C), \quad \beta_i \in H_1(C, \partial C), \quad i = 1, \ldots, r \tag{12}$$

as illustrated in Figure 3, where dotted line means that β_i continues on the other sheet. Note that $\alpha_i \cap \beta_j = \delta_{ij}$ and that

$$\overline{\alpha}_i = -\alpha_i, \quad \overline{\beta}_i = \beta_i, \tag{13}$$

where bar stands for complex conjugation. The ovals in Figure 2 represent the cycles α_i and $\beta_i - \beta_{i+1}$.

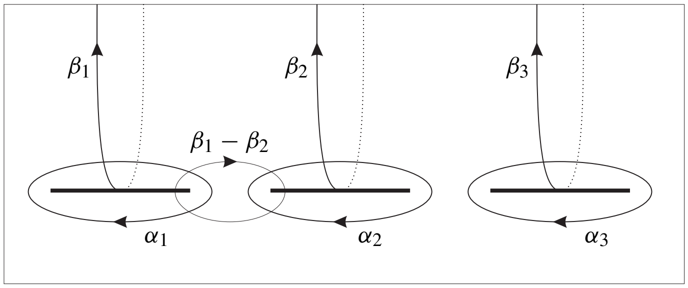


Figure 3. Cycles β_i go from $w = \infty$ to $w = 0$. Bold segments indicate bands.

The Seiberg–Witten differential

$$dS = \frac{1}{2\pi i}\, z\, \frac{dw}{w} = \pm \frac{r}{2\pi i}\left(1 + O\left(z^{-2}\right)\right)\, dz$$

is holomorphic except for a second order pole (without residue) at ∂C. Its derivatives with respect to $P \in \boldsymbol{U}$ are, therefore, holomorphic differentials on C. In fact, this gives

$$T_P \boldsymbol{U} \cong \text{holomorphic diff. on } C.$$

Nondegeneracy of periods implies the functions

$$a_i \overset{\text{def}}{=} \int_{\alpha_i} dS, \quad \sum a_i = 0, \tag{14}$$

which are real on M by (13), are local coordinates on $\boldsymbol{U}$, as are

$$a_i^\vee - a_{i+1}^\vee \overset{\text{def}}{=} 2\pi i \int_{\beta_i - \beta_{i+1}} dS, \quad \sum a_i^\vee = 0. \tag{15}$$

Further, there exists a function $\mathcal{F}(a; \Lambda)$, which is real and convex on M, such that

$$\left(\frac{\partial}{\partial a_i} - \frac{\partial}{\partial a_{i+1}}\right) \mathcal{F} = -\left(a_i^\vee - a_{i+1}^\vee\right). \tag{16}$$

Indeed, the Hessian of $\mathcal{F}$ equals $(-2\pi i)$ times the period matrix of C, hence symmetric (and positive definite on M). The function $\mathcal{F}$ is called the *Seiberg–Witten prepotential*. Note that $\mathcal{F}$ is multivalued on $\boldsymbol{U}$ and, in fact, its monodromy played a key role in the argument of Seiberg and Witten. By contrast, M is simply-connected, indeed

$$a^\vee : \mathsf{M} \to \mathsf{C}_+$$

is a diffeomorphism, see e.g. [18] for a more general result. Note that the periods (15) are the areas enclosed by the images of real ovals of C under $(z, w) \mapsto (z, \ln|w|)$. A similar geometric interpretation of the a_i's will be given in (40) below. In particular, the range

$$\mathsf{A} = a(\mathsf{M})$$

of the coordinates (14) is a proper subset of $\mathsf{C}_- = -\mathsf{C}_+$. At infinity of M, we have

$$a_i \sim \{\text{roots of } P\}, \quad a_1 \ll a_2 \ll \cdots \ll a_r .$$

2.5. Main result. We have now defined all necessary ingredients to confirm Nekrasov's conjecture in the following strong form:

Theorem 1 ([29]). *For $a \in \mathsf{A}$,*

$$-\lim_{\varepsilon \to 0} \varepsilon^2 \ln Z(\varepsilon; a; \Lambda) = \mathcal{F}(a; \Lambda), \tag{17}$$

where $\mathcal{F}$ is the Seiberg–Witten prepotential (16).

At the boundary of A, free energy has a singularity of the form

$$\mathcal{F} = -\left(a_i^\vee - a_j^\vee\right)^2 \ln\left(a_i^\vee - a_j^\vee\right) + \cdots$$

where dots denote analytic terms. This singularity is one of the main physical features of the Seiberg–Witten theory.

In broad strokes, the logic of the proof was explained in the Introduction. We now proceed with the details.

3. The random partition problem

3.1. Fixed points contributions. A rank 1 torsion-free sheaf on $\mathbb{C}^2$ is a fancy name to call an ideal I of $\mathbb{C}[x, y]$. Any partition λ defines one by

$$I_\lambda = (x^{\lambda_1}, x^{\lambda_2} y, x^{\lambda_3} y^2, \dots) \subset \mathbb{C}[x, y].$$

It is easy to see that all torus-fixed points of $\overline{\mathcal{M}}(r, n)$ have the form

$$\mathfrak{F} = \bigoplus_{k=1}^{r} I_{\lambda^{(k)}}, \quad \sum \left|\lambda^{(k)}\right| = n, \tag{18}$$

where $\lambda^{(k)}$ is an r-tuple of partitions. Our goal now is to compute the character of the torus action in the tangent space to the fixed point (18) and thus the contribution of $\mathfrak{F}$ to the sum in (6).

By construction of $\overline{\mathcal{M}}(r, n)$, its tangent space at $\mathfrak{F}$ equals $\mathrm{Ext}^1_{\mathbb{P}^2}(\mathfrak{F}, \mathfrak{F}(-L_\infty))$. From the vanishing of the other Ext-groups we conclude

$$\operatorname{tr} e^t\big|_{\mathrm{Ext}^1_{\mathbb{P}^2}(\mathfrak{F},\mathfrak{F}(-L_\infty))} = \mathcal{X}_{\mathcal{O}^{\oplus r}}(t) - \mathcal{X}_{\mathfrak{F}}(t), \tag{19}$$

where $\mathcal{X}_{\mathfrak{F}}(t)$ is the character

$$\mathcal{X}_{\mathfrak{F}}(t) = \operatorname{tr} e^t\big|_{\chi_{\mathbb{C}^2}(\mathfrak{F},\mathfrak{F})}$$

of the infinite-dimensional virtual representation

$$\chi_{\mathbb{C}^2}(\mathfrak{F},\mathfrak{F}) = \mathrm{Ext}^0_{\mathbb{C}^2}(\mathfrak{F},\mathfrak{F}) - \mathrm{Ext}^1_{\mathbb{C}^2}(\mathfrak{F},\mathfrak{F}) + \mathrm{Ext}^2_{\mathbb{C}^2}(\mathfrak{F},\mathfrak{F}).$$

Any graded free resolution of $\mathfrak{F}$ gives

$$\mathcal{X}_{\mathfrak{F}}(t) = |\mathsf{G}_{\mathfrak{F}}(t)|^2, \quad t \in \mathrm{Lie}(K),$$

where $\mathsf{G}_{\mathfrak{F}}(t)$ is, up to a factor, the character of $\mathfrak{F}$ itself

$$\begin{aligned}\mathsf{G}_{\lambda^{(1)},\dots,\lambda^{(r)}}(t) &= (e^{-i\varepsilon/2} - e^{i\varepsilon/2}) \operatorname{tr} e^t|_{\mathfrak{F}} \\ &= \sum_{k=1}^{r} e^{ia_k} \sum_{j=1}^{\infty} \exp\bigl(i\varepsilon\bigl(\lambda_j^{(k)} - j + \tfrac{1}{2}\bigr)\bigr).\end{aligned} \tag{20}$$

It is also a natural generating function of the r-tuple $\lambda^{(k)}$.

Note that the weight of any $\mathfrak{F}$ is real and positive, being a product of purely imaginary numbers in conjugate pairs.

3.2. Perturbative factor. In the spirit of the original uncompactified gauge theory problem on $\mathbb{R}^4$, we would like to drop the first term in (19) and declare its contribution canceled by Z_{pert}. In view of (20), this requires a regularization of the following product

$$Z_{\mathrm{pert}} \text{“=”} \prod_{k,k'=1}^{r} \prod_{j,j'=1}^{\infty} i(a_k - a_{k'} + \varepsilon(j-j')).$$

A natural regularization is provided by Barnes' double Γ-function (21), see e.g. [37]. For $c_1, c_2 \in \mathbb{R}$ and $\Re w \gg 0$, define

$$\zeta_2(s; w \mid c_1, c_2) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{dt}{t}\, t^s \, \frac{e^{-wt}}{\prod(1-e^{-c_i t})}.$$

This has a meromorphic continuation in s with poles at $s = 1, 2$. Define

$$\Gamma_2(w \mid c_1, c_2) = \exp \frac{d}{ds} \zeta(s; w \mid c_1, c_2)\big|_{s=0}. \tag{21}$$

Through the difference equation

$$w\, \Gamma_2(w)\, \Gamma_2(w + c_1 + c_2) = \Gamma_2(w + c_1)\, \Gamma_2(w + c_2) \tag{22}$$

it extends to a meromorphic function of w. We define

$$Z_{\mathrm{pert}} = \prod_{k,k'} \Gamma_2\left(\frac{i(a_k - a_{k'})}{\Lambda} \,\middle|\, \frac{i\varepsilon}{\Lambda}, \frac{-i\varepsilon}{\Lambda}\right)^{-1}. \tag{23}$$

where Γ_2 is analytically continued to imaginary arguments using

$$\Gamma_2\,(Mw\,|\,Mc,-Mc) = M^{\frac{w^2}{2c^2}-\frac{1}{12}}\;\Gamma_2(w\,|\,c,-c), \qquad M \notin (-\infty,0].$$

The scaling by Λ is introduced in (23) to make (8) homogeneous of degree 0 in a, ε, and Λ. Note also

$$\Gamma_2(0|\,1,-1) = e^{-\zeta'(-1)}.$$

Our renormalization rule (23) fits nicely with the following transformation of the partition function Z.

3.3. Dual partition function. For $r = 1$, the weight of I_λ in (8) equals

$$\Lambda^{2n}\ \det{}^{-1} t\big|_{T_{I_\lambda}\bar{\mathcal{M}}(1,n)} = \frac{1}{n!}\left(\frac{\Lambda^2}{\varepsilon^2}\right)^n \mathfrak{M}_{\mathrm{Planch}}(\lambda), \tag{24}$$

where $\mathfrak{M}_{\mathrm{Planch}}$ is the Plancherel measure (1) and the prefactor is the Poisson weight with parameter Λ^2/ε^2. For $r > 1$, we will transform Z into the partition function (29) of the Plancherel measure in a *periodic potential* with period r.

Let a function $\xi : \mathbb{Z}+\frac{1}{2} \to \mathbb{R}$ be periodic with period r and mean 0. The energy $\Xi(\lambda)$ of the configuration $\mathfrak{S}(\lambda)$ in the potential ξ is defined by Abel's rule

$$\Xi(\lambda) = \sum_{x\in\mathfrak{S}(\lambda)} \xi(x) \stackrel{\mathrm{def}}{=} \lim_{z\to+0} \sum_{x\in\mathfrak{S}(\lambda)} \xi(x)\, e^{zx}.$$

Grouping the points of $\mathfrak{S}(\lambda)$ modulo r uniquely determines an r-tuple of partitions $\lambda^{(k)}$, known as r-quotients of λ, and shifts $s_k \in \mathbb{Q}$ such that

$$\mathfrak{S}(\lambda) = \bigsqcup_{k=1}^{r} r\big(\mathfrak{S}\big(\lambda^{(k)}\big) + s_k\big) \tag{25}$$

and

$$rs \equiv \rho \mod r\,\mathbb{Z}_0^r, \qquad \rho = \left(\tfrac{r-1}{2},\ldots,\tfrac{1-r}{2}\right),$$

where $\mathbb{Z}_0^r$ denotes vectors with zero sum. It follows from (25) that

$$\mathsf{G}_\lambda\,(\varepsilon/r) = \mathsf{G}_{\lambda^{(1)},\ldots,\lambda^{(r)}}(\varepsilon;\,\varepsilon s). \tag{26}$$

Letting $\varepsilon \to 2\pi i k$, $k = 1,\ldots,r-1$, in (26) gives

$$\Xi(\lambda) = (s,\xi) = \sum s_i\,\xi_i, \quad \xi_i = \xi\left(\tfrac{1}{2}-i\right), \tag{27}$$

while the $\varepsilon \to 0$ limit in (26) yields

$$|\lambda| = r\left(\sum\big|\lambda^{(k)}\big| + \sum \frac{s_k^2}{2}\right) + \frac{1-r^2}{24}.$$

Using these formulas and the difference equation (22), we compute

$$Z^\vee(\varepsilon;\xi_1,\dots,\xi_r;\Lambda) \overset{\text{def}}{=} \sum_{a\in\varepsilon(\rho+r\mathbb{Z}_0^r)} \exp\left(\frac{(\xi,a)}{r\varepsilon^2}\right) Z(r\varepsilon;a;\Lambda) \tag{28}$$

$$= e^{\zeta'(-1)+\frac{\pi i}{24}} \sum_\lambda \left|\frac{\Lambda}{\varepsilon}\right|^{2|\lambda|-\frac{1}{12}} \left(\frac{\dim\lambda}{|\lambda|!}\right)^2 \exp\left(\frac{\Xi(\lambda)}{\varepsilon}\right). \tag{29}$$

We call (28) the *dual partition function*. By (29), it equals the partition function of a periodically weighted Plancherel measure on partitions.

While it will play no role in what follows, it may be mentioned here that $Z^\vee$ is a very interesting object to study not asymptotically but exactly. For example, Toda equation for $\ln \mathbb{Z}^\vee$ may be found in Section 5 of [29].

3.4. Dual free energy. Define the dual free energy by

$$\mathcal{F}^\vee(\xi;\Lambda) = -\lim_{\varepsilon\to 0} \varepsilon^2 \ln Z^\vee. \tag{30}$$

Since (28) is a Riemann sum for Laplace transform, we may expect that

$$\mathcal{F}^\vee(\xi;\Lambda) = \min_{a\in\mathbb{R}_0^r} \frac{1}{r^2}\,\mathcal{F}(a;\Lambda) - \frac{1}{r}\,(\xi,a) \tag{31}$$

that is, up to normalization, $\mathcal{F}^\vee$ is the Legendre transform of $\mathcal{F}$. This is because the asymptotics of Laplace transform is determined by one point – the maximum. Our plan is apply to same logic to the infinite-dimensional sum (29), namely, to show that its $\varepsilon\to 0$ asymptotics is determined by a single term, the *limit shape*.

The law of large numbers, a basic principle of probability, implies that on a large scale most random system are deterministic: solids have definite shape, fluids obey the laws of hydrodynamics, etc. Only magnification reveals the full randomness of nature.

In the case at hand, the weight of a partition λ in (29), normalized by the whole sum, defines a probability measure on the set of partitions. This measure depends on a parameter ε and as $\varepsilon\to 0$ it clearly favors partitions of larger and larger size. In fact, the expected size of λ grows as ε^{-2}. We thus expect the diagram of λ, scaled by ε in both directions, to satisfy a law of large numbers, namely, to have a nonrandom limit shape. By definition, this limit shape will dominate the leading $\varepsilon\to 0$ asymptotics of $Z^\vee$. In absence of the periodic potential Ξ, such analysis is a classical result of Logan–Shepp and Vershik–Kerov [21], [42], [43].

Note that the maximum in (31) is over all of a, including the problematic region where $|a_i - a_j|$ get small. However, this region does not contribute to $\mathcal{F}^\vee$ as the convexity of free energy is lost there. We will see this reflected in the following properties of $\mathcal{F}^\vee$: it is strictly concave, analytic in the interior of the Weyl chambers, and singular along the chambers' walls.

3.5. Variational problem for the limit shape. The *profile* of a partition λ is, by definition, the piecewise linear function plotted in bold in Figure 1. Let ψ_λ be the profile of λ scaled by ε in both directions. The map $\lambda \mapsto \psi_\lambda$ embeds partitions into the convex set Ψ of functions ψ on $\mathbb{R}$ with Lipschitz constant 1 and

$$|\psi| = \int |\psi(x) - |x|| \, dx < \infty.$$

The Lipschitz condition implies

$$\|\psi_1 - \psi_2\|_C \leq \|\psi_1 - \psi_2\|_{L^1}^{1/2}, \quad \psi_1, \psi_2 \in \Psi,$$

and so Ψ is complete and separable in the L^1-metric. Some function of a partition have a natural continuous extension to Ψ, for example

$$\mathsf{G}_\lambda(\varepsilon) = \frac{1}{e^{i\varepsilon/2} - e^{-i\varepsilon/2}} \left(1 - \frac{1}{2} \int e^{ix} \, (\psi_\lambda(x) - |x|) \, dx \right),$$

while others, specifically the ones appearing in (29), do not. An adequate language for dealing with this is the following.

Let $f(\lambda) \geq 0$ be a function on partitions depending on the parameter ε. We say that it satisfies a *large deviation* principle with action (rate) functional $\mathcal{S}_f(\psi)$ if for any set $A \subset \Psi$

$$-\lim \varepsilon^2 \ln \sum_{\psi_\lambda \in A} f(\lambda) \subset \left[\inf_{\bar{A}} \mathcal{S}_f, \inf_{A^\circ} \mathcal{S}_f \right] \subset \mathbb{R} \cup \{+\infty\}, \tag{32}$$

where lim denotes all limit points, A° and $\bar{A}$ stand for the interior and closure of A, respectively.

For the Plancherel weight (24), Logan–Shepp and Vershik–Kerov proved a large deviation principle with action

$$\mathcal{S}_{\mathrm{pl}}(\psi) = \frac{1}{2} \int_{x<y} (1 + \psi'(x))(1 - \psi'(y)) \ln \frac{|x-y|}{\Lambda} \, dx \, dy. \tag{33}$$

Note that in this case the sum in (32) may be replaced by maximum because the number of partitions of n grows subexponentially in n. In other words, there is no entropic contribution in (33).

The periodic potential $\Xi(\lambda)$ produces a *surface tension* addition to the total action $\mathcal{S}$

$$\mathcal{S} = \mathcal{S}_{\mathrm{pl}} + \mathcal{S}_{\mathrm{surf}}, \quad \mathcal{S}_{\mathrm{surf}}(\psi) = \frac{1}{2} \int \sigma(\psi') \, dt,$$

where σ is a convex piecewise-linear function of the kind plotted in Figure 4. It is linear on segments of length $2/r$ with slopes $\{\xi_i\}$, in increasing order. The form of $\mathcal{S}_{\mathrm{surf}}$ is easy to deduce directly; it can also be seen as e.g. the most degenerate case of

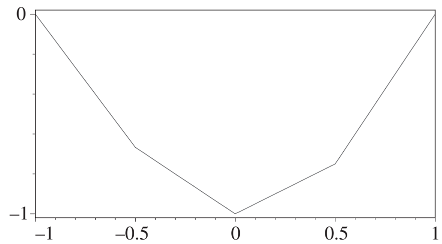


Figure 4. The surface tension σ for $r = 4$ and $\xi = \left\{-\frac{4}{3}, -\frac{2}{3}, \frac{1}{2}, \frac{3}{2}\right\}$.

the surface tension formula from Section 4.2. The singularities in the surface tension σ are responsible for *facets*, that is, linear pieces, in the minimizer, see Figure 6. The slopes of these facets are precisely the points where σ' is discontinuous[1]. Note that σ and hence $\mathcal{S}$ is a symmetric function of the ξ_i's.

The functional $\mathcal{S}$ is strictly convex and its sublevel sets $\mathcal{S}^{-1}((-\infty, c])$ are compact, which can be seen by rewriting $\mathcal{S}_{\mathrm{pl}}$ in terms of the Sobolev $H^{1/2}$ norm, see [21], [43]. Therefore it has a unique minimum $\psi_\star$ – the limit shape. The large deviation principle and the definition of the dual free energy and (30) together imply

$$\mathcal{F}^{\vee}(\xi; \Lambda) = \mathcal{S}(\psi_\star). \tag{34}$$

Our business, therefore, is to find this minimizer $\psi_\star$.

3.6. The minimizer. By convexity, a local minimum of $\mathcal{S}$ is automatically a global one. Since σ has one-sided derivatives, a local minimum can be characterized by nonnegativity of all directional derivatives. This leads to the following *complementary slackness* conditions for the convolution of $\psi_\star''(x)$ with the kernel

$$\mathsf{L}(x) = x \ln \frac{|x|}{\Lambda} - x = \int_0^x \ln \left|\frac{y}{\Lambda}\right| \, dy.$$

There exists a constant c_0, which is the Lagrange multiplier from the constraint $\int \delta\psi' = 0$, such that

$$\begin{aligned} &\mathsf{L} * \psi_\star''(x) + c_0 = \xi_i, && \psi_\star'(x) \in \left(-1 + \tfrac{2i-2}{r}, -1 + \tfrac{2i}{r}\right), \\ &\mathsf{L} * \psi_\star''(x) + c_0 \in [\xi_i, \xi_{i+1}], && \psi_\star'(x) = -1 + \tfrac{2i}{r}, \end{aligned} \tag{35}$$

where to simplify notation we assumed that

$$\xi \in \mathsf{C}_-, \quad \xi_0 = -\infty, \quad \xi_{r+1} = +\infty.$$

[1]There are many advantages in viewing random partitions as 2-dimensional slices of random 3-dimensional objects discussed in Section 4. From the probability viewpoint, this links random partitions with rather realistic models of crystalline surfaces with local interaction, enriching both techniques and intuition. In particular, coexistence of facets and curved regions in our limit shapes is the same phenomenon as observed in natural crystals. From the gauge theory viewpoint, it is also very natural, especially in the context of 5-dimensional theory on $\mathbb{R}^4 \times S^1$, which corresponds to the K-theory of the instanton moduli spaces.

Recall that C_- denotes the negative Weyl chamber.

The function $\psi_\star''$ will turn out to be nonnegative and supported on a union of r intervals, which are precisely the *bands* of Section 2.4. The gaps will produce the facets in the limit shape.

It is elementary to see that for a maximal curve (11) the map

$$\Phi(z) = 1 + \frac{2}{\pi i r} \ln w = 1 + \frac{2}{\pi i} \ln \frac{\Lambda}{z} + O(z^{-1}), \quad z \to \infty \tag{36}$$

where w is the smaller root of (11), defines a conformal map of the upper half-plane to a slit half-strip

$$\Delta \subset \{z \,|\, \Im z > 0, |\Re z| < 1\}$$

as in Figure 5. The slits in Δ go along

$$\Re z = -1 + 2i/r, \quad i = 1, \dots, r-1,$$

and their lengths are, essentially, the critical values of the polynomial $P(z)$. The bands and gaps are preimages of the horizontal and vertical segments of $\partial\Delta$, respectively.

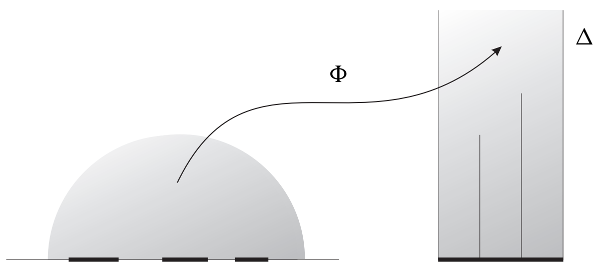


Figure 5. Conformal map defined by a maximal curve.

We claim that

$$\psi_\star' = \Re\Phi\big|_{\mathbb{R}}, \tag{37}$$

where the polynomial $P(z)$ is determined by the relation (39) below. The equations (35) are verified for (37) as follows. Since $\Phi'(z) = O(z^{-1})$, $z \to \infty$, we have the Hilbert transform relation

$$\text{P.V.}\,\frac{1}{x} * \Re\Phi'\big|_{\mathbb{R}} = \pi \Im\Phi'\big|_{\mathbb{R}}.$$

Integrating it once and using (36) to fix the integration constant, we get

$$(\mathsf{L} * \Re\Phi')' = \pi \Im\Phi.$$

Therefore, the function $\mathsf{L} * \Re\Phi'$ is constant on the bands and strictly increasing on the gaps, hence (37) satisfies (35) with

$$\xi_{i+1} - \xi_i = \pi \int_{i\text{th gap}} \Im\Phi(x)\, dx \tag{38}$$

Integrating (38) by parts and using definitions from Section 2.4 gives

$$\xi = -\frac{a^{\vee}}{r}, \tag{39}$$

thus every limit shape $\psi_\star$ comes from a maximal curve. For example, the limit shape corresponding to the curve from Figure 2 is plotted in Figure 6. Note also that for $C \in \mathsf{M}$, we have

$$a_i = \frac{r}{2}\left(I_{i-1} - I_i\right), \tag{40}$$

where I_i is the intercept of the ith facet of the limit shape. In particular, $\mathsf{A} \subset \mathsf{C}_-$.

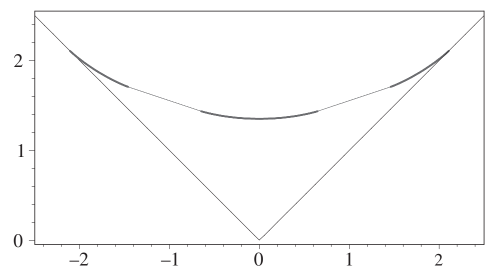


Figure 6. Limit shape corresponding to the curve from Figure 2. Thin segments are facets.

For given $\xi \in \mathsf{C}_-$, consider the distribution of the r-quotients $\lambda^{(i)}$ of the partition λ, as defined in Section 3.3. For the shifts s_k in (25) we have using (27)

$$\varepsilon s \to -\frac{\partial \mathcal{F}^{\vee}}{\partial \xi}, \quad \varepsilon \to 0,$$

in probability. Observe that

$$\frac{\partial}{\partial \xi}\mathcal{F}^{\vee}(\xi) = \left[\frac{\partial}{\partial \xi}\mathcal{S}\right](\psi_\star)$$

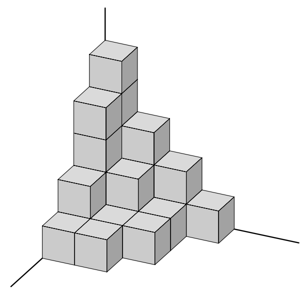
since the other term, containing $\frac{\partial}{\partial \xi}\psi_\star$, vanishes by the definition of a maximum. Definitions and integration by parts yield

$$-\left(\frac{\partial}{\partial \xi_i} - \frac{\partial}{\partial \xi_{i+1}}\right)\mathcal{F}^{\vee} = \frac{a_i - a_{i+1}}{r}.$$

By (26), this means that the resulting sum over the r-quotients $\lambda^{(i)}$ is the original partition function Z with parameters $a \in \mathsf{A}$. This concludes the proof.

4. The next dimension

4.1. Stepped surfaces



An obvious 3-dimensional generalization of a partition, also known as a plane partition can be seen on the right. More generally, we consider *stepped surfaces*, that is, continuous surfaces glued out of sides of a unit cube, spanning a given polygonal contour in $\mathbb{R}^3$, and projecting 1-to-1 in the (1, 1, 1) direction, see Figure 7. Note that stepped surfaces minimize the surface area for given boundary conditions, hence can be viewed as zero temperature limit of the interface in the 3D Ising model.

The most natural measure on stepped surfaces is the uniform one with given boundary conditions, possibly conditioned on the volume enclosed. It induces Plancherel-like measures on 2-dimensional slices. Stepped surfaces are in a natural bijection with fully packed dimers on the hexagonal lattice and Kasteleyn theory of planar dimers [16] forms the basis of most subsequent developments.

The following law of large numbers for stepped surfaces was proven in [7]. Let C_n be a sequence of boundary contours such that each C_n can be spanned by at least one stepped surface. Suppose that $n^{-1}C_n$ converge to a given curve $C \subset \mathbb{R}^3$. Then, scaled by n^{-1}, uniform measures on stepped surfaces spanning C_n converge to the δ-measure on a single Lipschitz surface spanning C – the limit shape. This limit shape formation is clearly visible in Figure 7.

Figure 7. A limit shape simulation. The frozen boundary is the inscribed cardioid.

The limit shape is the unique minimizer of the following functional. Let the surface be parameterized by $x_3 = h(x_3 - x_1, x_3 - x_2)$, where h is a Lipschitz function with gradient in the triangle $\triangle$ with vertices $(0,0)$, $(0,1)$, $(1,0)$. Let Ω be the planar region enclosed by the projection of C in the $(1,1,1)$ direction. We will use $(x,y) = (x_3 - x_1, x_2 - x_1)$ as coordinates on Ω. The limit shape is the unique minimizer of

$$\mathcal{S}_{\mathrm{step}}(h) = \int_\Omega \sigma_{\mathrm{step}}(\nabla h)\, dx\, dy, \tag{41}$$

where, in the language of [20], the surface tension σ_{step} is the Legendre dual of the Ronkin function of the straight line

$$z + w = 1. \tag{42}$$

We recall that for a plane curve $P(z,w) = 0$, its Ronkin function [23] is defined by

$$R(x,y) = \frac{1}{(2\pi i)^2} \iint_{\substack{|z|=e^x \\ |w|=e^y}} \log|P(z,w)|\, \frac{dz}{z}\, \frac{dw}{w}. \tag{43}$$

The gradient ∇R always takes values in the Newton polygon $\Delta(P)$ of the polynomial P, so $\Delta(P)$ is naturally the domain of the Legendre transform $R^\vee$. For the straight line as above, the Newton polygon is evidently the triangle $\triangle$.

The surface tension σ_{step} is singular and not strictly convex at the boundary of $\triangle$, which leads to formation of *facets* and *edges* in the limit shape (which can be clearly seen in Figure 7). This models facet formation in natural interfaces, e.g. crystalline surfaces, and is the most interesting aspect of the model. Note that facets are completely ordered (or *frozen*). The boundary between the ordered and disordered (or *liquid*) regions is known as the *frozen boundary*.

The following transformation of the Euler-Lagrange equation for (41) found in [19] greatly facilitates the study of the facet formation. Namely, in the liquid region we have

$$\nabla h = \frac{1}{\pi}(\arg w, -\arg z), \tag{44}$$

where the functions z and w solve the differential equation

$$\frac{z_x}{z} + \frac{w_y}{w} = c \tag{45}$$

and the algebraic equation (42). Here c is the Lagrange multiplier for the volume constraint $\int_\Omega h = \mathrm{const}$, the unconstrained case is $c = 0$. At the boundary of the liquid region, z and w become real and the ∇h starts to point in one of the coordinate directions.

The first-order quasilinear equation (45) is, essentially, the complex Burgers equation $z_x = z z_y$ and, in particular, it can be solved by complex characteristics as follows. There exists an analytic function $Q(z,w)$ such that

$$Q(e^{-cx} z, e^{-cy} w) = 0. \tag{46}$$

In other words, $z(x, y)$ can be found by solving (42) and (46). In spirit, this is very close to Weierstraß parametrization of minimal surfaces in terms of analytic data.

Frozen boundary can only develop if Q is real, in which case the roots (z, w) and $(\bar{z}, \bar{w})$ of (46) coincide at the frozen boundary. At a smooth point of the frozen boundary, the multiplicity of this root will be exactly two, hence ∇h has a square-root singularity there. As a result, the limit shape has an $x^{3/2}$ singularity at the generic point of the frozen boundary, thus recovering the well-known Pokrovsky–Talapov law [36] in this situation. At special points of the frozen boundary, triple solutions of (46) occur, leading to a cusp singularity. One such point can be seen in Figure 7.

Remarkably, for a dense set of boundary condition the function Q is, in fact, a polynomial. Consequently, the frozen boundary takes the form $R(e^{cx}, e^{cy}) = 0$, where R is the polynomial defining the planar dual of the curve $Q = 0$. This allows to use powerful tools of algebraic geometry to study the singularities of the solutions, see [19]. The precise result proven there is

Theorem 2 ([19]). *Suppose the boundary contour C is a connected polygon with $3k$ sides in coordinate directions (cyclically repeated) which can be spanned by a Lipschitz function with gradient in $\triangle$. Then $Q = 0$ is an algebraic curve of degree k and genus zero.*

For example, for the boundary contour in Figure 7 we have $k = 3$ (one of the boundary edges there has zero length) and hence R is the dual of a degree 3 genus 0 curve – a cardioid. The procedure of determining Q from the boundary conditions is effective and can be turned into a practical numeric homotopy procedure, see [19]. Higher genus frozen boundaries occur for multiply-connected domains, in fact, the genus of Q equals the genus of the liquid region.

Of course, for a probabilist, the law of large numbers is only the beginning and the questions about CLT corrections to the limit shape and local statistics of the surface in various regions of the limit shape follow immediately. Conjecturally, the limit shape controls the answers to all these questions. For example, the function $e^{-cx}z$ defines a *complex structure* on the liquid region and, conjecturally, the Gaussian correction to the limit shape is given by the *massless free field* in the corresponding conformal structure. In the absence of frozen boundaries and without the volume constraint, this is proven in [17]. See e.g. [15], [17], [20], [32] for an introduction to the local statistics questions.

4.2. Periodic weights. Having discussed periodically weighted Plancherel measure and a 3-dimensional analog of the Plancherel measure, we now turn to periodically weighted stepped surfaces. This is very natural if stepped surfaces are interpreted as crystalline interfaces. Periodic weights are introduced as follows: we weight each square by a periodic function of $x_3 - x_1$ and $x_2 - x_1$ (with some integer period M).

The role previously played by the straight line (42) is now played by a certain higher degree curve $P(z, w) = 0$, the spectral curve of the corresponding periodic

Kasteleyn operator. In particular, the surface tension σ_{step} is now replaced by the Legendre dual of the Ronkin function of P, see [20]. We have

$$\deg P = M$$

and the coefficients of P depend polynomially on the weights.

The main result of [20], known as *maximality*, says that for real and positive weights the curve P is always a real algebraic curve of a very special kind, namely, a *Harnack curve*, see [23]. Conversely, as shown in [18], all Harnack curves arise in this way.

Harnack curves are, in some sense, the best possible real curves; their many remarkable properties are discussed in [23]. One of several equivalent definitions of a Harnack curve is that the map

$$(z, w) \mapsto (\log |z|, \log |w|) \tag{47}$$

from $P(z, w) = 0$ to $\mathbb{R}^2$ is 1-to-1 on the real locus of P and 2-to-1 over the rest. The image of $P = 0$ under (47) is known as the *amoeba* of P. Note from (43) that the gradient ∇R of the Ronkin function of P is nonconstant precisely for (x, y) in the amoeba of P. In other words, the Ronkin function has a facet (that is, a linear piece) over every component of the amoeba complement. The 2-to-1 property implies that the number of compact facets of Ronkin function equals the (geometric) genus of the curve P. Each of these facets translates into the singularity of the surface tension and, hence, into facets with the same slope in limit shapes.

By Wulff's theorem, the Ronkin function itself is a minimizer, corresponding to its own ("crystal corner") boundary conditions. An example of the Ronkin function of a genus 1 Harnack curve can be seen in Figure 8.

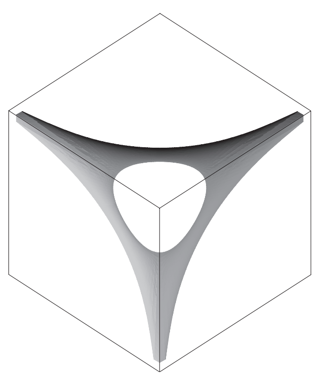


Figure 8. The (curved part of the) Ronkin function of a genus 1 curve. Its projection to the plane is the amoeba.

Maximality implies *persistence of facets*, namely, for fixed period M, there will be $\binom{M-1}{2}$ compact facets of the Ronkin function and $\binom{M-1}{2}$ corresponding singularities of the surface tension, except on a codimension 2 subvariety of the space of weights. It also implies e.g. the following *universality of height fluctuations* in the liquid region

$$\operatorname{Var}(h(a)-h(b)) \sim \frac{1}{\pi} \ln \|a-b\|, \quad \|a-b\| \to \infty.$$

Remarkably, formulas (44), (45), and (46) need no modifications for periodic weights. Replacing (42) by $P(z,w)=0$ is the only change required, see [19].

From our experience with periodically weighted Plancherel measure, it is natural to expect that, for some special boundary conditions, the partition function of periodically weighted stepped surfaces will encode valuable physical information. A natural choice of "special boundary conditions" are the those of a *crystal corner*, when we require the surface to be asymptotic to given planes at infinity, as in Figure 8. For convergence of the partition function, one introduces a fugacity factor q^{vol}, where the missing volume is measured with respect to the "full corner".

I hope that further study will reveal many special properties of such crystal corner partition functions. Their extremely degenerate limits have been identified with all-genera, all-degree generating functions for Donaldson–Thomas invariants of toric Calabi–Yau threefolds. Namely, as the periodic weights become extreme, all limit shapes, and the Ronkin function in particular, degenerate to piecewise linear functions. This is known as the *tropical limit*. The only remaining features of limit shapes are the edges and the triple points, where 2 and 3 facets meet, respectively. In this tropical limit, the partition function becomes the partition function of ordinary, unweighted, 3D partitions located at triple points. These 3D partitions may have infinite legs along the edges, as in Figure 9 and through these legs they interact with their neighbors. This description precisely matches the localization formula for Donaldson–Thomas invariants of the toric threefold whose toric polyhedron is given by the piecewise linear limit shape, see [22].

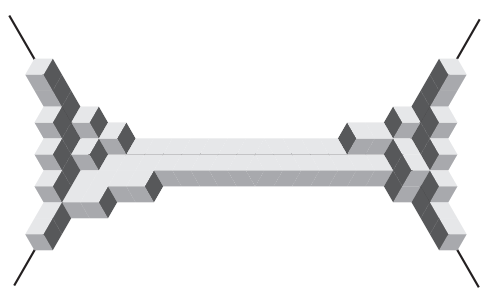


Figure 9. Two 3D partitions connected at an angle through an infinite leg.

Donaldson–Thomas theory of any 3-fold has been conjectured to be equivalent, in a nontrivial way, to the Gromov–Witten theory of the same 3-fold in [22]. For the toric Calabi–Yau 3-folds, this specializes to the earlier *topological vertex* conjecture of [1]. It is impossible to adequately review this subject here, see [35] for an introduction.

This is also related to the supersymmetric gauge theories considered in Section 2, or rather their 5-dimensional generalizations, via a procedure called *geometric engineering* of gauge theories. See for example [13] and references therein.

I find such close and unexpected interaction between rather basic statistical models and instantons in supersymmetric gauge and string theories very exciting and promising. The field is still full of wide open questions and, in my opinion, it is also full of new phenomena waiting to be discovered.

References

[1] Aganagic, M., Klemm, A., Marino, M., Vafa, C., The Topological Vertex. *Comm. Math. Phys.* **254** (2005), 425–478.

[2] Atyiah, M., Bott, R., The moment map and equivariant cohomology. *Topology* **23** (1) (1984), 1–28.

[3] Baik, J., Deift, P., Johansson, K., On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** (1999), 1119–1178.

[4] Borodin, A., Okounkov, A., Olshanski, G., Asymptotics of the Plancherel measures for symmetric groups. *J. Amer. Math. Soc.* **13** (3) (2000), 481–515.

[5] Braverman, A., Instanton counting via affine Lie algebras. I. Equivariant J-functions of (affine) flag manifolds and Whittaker vectors. In *Algebraic structures and moduli spaces*, CRM Proc. Lecture Notes 38, Amer. Math. Soc., Providence, RI, 2004, 113–132.

[6] Braverman, A., Etingof, P., Instanton counting via affine Lie algebras II: from Whittaker vectors to the Seiberg-Witten prepotential. math.AG/0409441.

[7] Cohn, H., Kenyon, R., Propp, J., A variational principle for domino tilings. *J. Amer. Math. Soc.* **14** (2) (2001), 297–346.

[8] D'Hoker, E., Phong, D., Lectures on supersymmetric Yang-Mills theory and integrable systems. In *Theoretical physics at the end of the twentieth century*, CRM Ser. Math. Phys., Springer-Verlag, New York 2002, 1–125.

[9] Donaldson, S, Kronheimer, P., *The geometry of four-manifolds*. Oxford Math. Monogr., The Clarendon Press, New York 1990.

[10] Dorey, N., Hollowood, T., Khoze, V., Mattis, M., The calculus of many instantons. *Phys. Rep.* **371** (4–5) (2002), 231–459.

[11] Hollowood, T., Iqbal, A., Vafa, C., Matrix models, geometric engineering, and elliptic genera. hep-th/0310272.

[12] Huybrechts, D., Lehn, M., *The geometry of moduli spaces of sheaves*. Aspects Math. E31, Vieweg, Braunschweig 1997.

[13] Iqbal,A., Kashani-Poor, A.-K., The vertex on a strip. hep-th/0410174.

[14] Johansson, K., Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. of Math.* **153** (1) (2001), 259–296.

[15] Johansson, K., Random matrices and determinantal processes. math-ph/0510038.

[16] Kasteleyn, P., Graph theory and crystal physics. In *Graph Theory and Theoretical Physics*, Academic Press, London 1967, 43–110.

[17] Kenyon, R., Height fluctuations in honeycomb dimers. math-ph/0405052.

[18] Kenyon, R., Okounkov, A., Planar dimers and Harnack curves. math.AG/0311062.

[19] Kenyon, R., Okounkov, A., Limit shapes and complex Burgers equation. math-ph/0507007.

[20] Kenyon, R., Okounkov, A., Sheffield, S., Dimers and amoebae. math-ph/0311005.

[21] Logan, B., Shepp, L., A variational problem for random Young tableaux. *Adv. Math.* **26** (1977), 206–222.

[22] Maulik, D., Nekrasov, N., Okounkov, A., Pandharipande, R., Gromov-Witten theory and Donaldson-Thomas theory, I. & II. math.AG/0312059, math.AG/0406092.

[23] Mikhalkin, G., Amoebas of algebraic varieties and tropical geometry. In *Different faces of geometry*, Int. Math. Ser. (N. Y.), Kluwer/Plenum, New York 2004, 257–300.

[24] Nakajima, H., *Lectures on Hilbert schemes of points on surfaces*. Univ. Lecture Ser. 18, Amer. Math. Soc., Providence, RI, 1999.

[25] Nakajima, H., Yoshioka, K., Instanton counting on blowup. I. 4-dimensional pure gauge theory. *Invent. Math.* **162** (2005), 313–355.

[26] Nakajima, H., Yoshioka, K., Lectures on instanton counting. *Algebraic structures and moduli spaces*, CRM Proc. Lecture Notes, 38, Amer. Math. Soc., Providence, RI, 2004, 31–101.

[27] Nakajima, H., Yoshioka, K., Instanton counting on blowup. II. K-theoretic partition function. math.AG/0505553.

[28] Nekrasov, N., Seiberg-Witten prepotential from instanton counting. *Adv. Theor. Math. Phys.* **7** (5) (2003), 831–864.

[29] Nekrasov, N., Okounkov, A., Seiberg-Witten Theory and Random Partitions. In *The Unity of Mathematics* (ed. by P. Etingof, V. Retakh, I. M. Singer), Progr. Math. 244, Birkhäuser, Boston, MA, 2006, 525–596.

[30] Nekrasov, N., Shadchin, S., ABCD of instantons. *Comm. Math. Phys.* **252** (2004), 359–391.

[31] Okounkov, A., Random matrices and random permutations. *Internat. Math. Res. Notices* **2000** (20) (2000), 1043–1095.

[32] Okounkov, A., Symmetric function and random partitions. In *Symmetric functions 2001: surveys of developments and perspectives* (ed. by S. Fomin), Kluwer Acad. Publ., Dordrecht 2002, 223–252.

[33] Okounkov, A., The uses of random partitions. math-ph/0309015.

[34] Okounkov, A., Reshetikhin, N., Vafa, C., Quantum Calabi-Yau and Classical Crystals. In *The Unity of Mathematics* (ed. by P. Etingof, V. Retakh, I. M. Singer), Progr. Math. 244, Birkhäuser, Boston, MA, 2006, 597–618.

[35] Okounkov, A., Random surfaces enumerating algebraic curves. In *Proceedings of Fourth European Congress of Mathematics*, EMS, Zürich 2005, 751–768.

[36] Pokrovsky, V., Talapov, A., Theory of two-dimensional incommensurate crystals. *Soviet Phys. JETP* **78** (1) (1980), 269–295.

[37] Ruijsenaars, S., On Barnes' multiple zeta and gamma functions. *Adv. Math.* **156** (1) (2000), 107–132.

[38] Seiberg, N., Witten, E., Electric-magnetic duality, monopole condensation, and confinement in $\mathcal{N} = 2$ supersymmetric Yang-Mills theory. *Nuclear Phys. B* **426** (1994), 19–52; Erratum ibid. **430** (1994), 485–486.

[39] Seiberg, N., Witten, E., Monopoles, duality and chiral symmetry breaking in $\mathcal{N} = 2$ supersymmetric QCD. *Nuclear Phys. B* **431** (1994), 484–550.

[40] Sodin, M., Yuditskii, P., Functions that deviate least from zero on closed subsets of the real axis. *St. Petersburg Math. J.* **4** (2) (1993), 201–249.

[41] Toda, M., *Theory of nonlinear lattices*. Springer Ser. Solid-State Sci. 20, Springer-Verlag, Berlin 1981.

[42] Vershik, A., Kerov, S., Asymptotics of the Plancherel measure of the symmetric group and the limit form of Young tableaux. *Soviet Math. Dokl.* **18** (1977), 527–531.

[43] Vershik, A., Kerov, S., Asymptotics of maximal and typical dimensions of irreducible representations of a symmetric group. *Funct. Anal. Appl.* **19** (1) (1985), 21–31 .

[44] Witten, E., Dynamics of quantum field theory. In *Quantum fields and strings: a course for mathematicians* (ed. by P. Deligne, P. Etingof, D. Freed, L. Jeffrey, D. Kazhdan, J. Morgan, D. Morrison and E. Witten), Vol. 2, Amer. Math. Soc., Providence, RI, IAS, Princeton, NJ, 1999, 1119–1424.

Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, New Jersey, 08544, U.S.A.

E-mail: okounkov@math.princeton.edu