

# The Kac-Rice formula: basic definitions and a first application

Antoine Maillard

February 18, 2020

These notes are based on a brief introduction given at the Kavli Institute for Theoretical Physics (Santa Barbara) in February 2019.

## 1 The Kac-Rice formula

### 1.1 The area formula

The Kac-Rice formula is intuitively not derived from any involved probabilistic tool, but rather is a consequence of a purely geometric result, called the area formula (itself a consequence of the more general coarea formula), described in much more generality by Federer [Fed59], and stated for instance in [AW09]. This formula is the generalization of the following non-rigorous intuition: for a smooth function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , and  $T \subseteq \mathbb{R}$ , denoting  $N_f(u, T)$  the number of solutions to the equation  $f(t) = u$  with  $t \in T$ , one would want to write informally:

$$N_f(u, T) = \int_{f(T)} \delta(v - u) dv, \tag{1}$$

$$= \int_T \delta(f(t) - u) |f'(t)| dt. \tag{2}$$

The area formula generalizes and makes rigorous this intuition, by showing a weak version of this last equality. We follow here the statement of [AW09].

**Proposition 1.1 (Area formula)** *Let  $f : U \rightarrow \mathbb{R}^d$  be a  $C^1$  function defined on an open subset  $U$  of  $\mathbb{R}^d$ . Assume that the sets of critical values of  $f$  has zero Lebesgue measure, and denote  $N_f(u, T)$  the number of solutions to the equation  $f(t) = u$  with  $t \in T$ . Then, for any Borel set  $T \subset \mathbb{R}^d$ , and any  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  continuous and bounded:*

$$\int_{\mathbb{R}^d} g(u) N_f(u, T) = \int_T |\det f'(t)| g(f(t)) dt. \tag{3}$$

Note that in a large part of the theoretical physics literature, the area formula (and subsequently the Kac-Rice formula) is written in the form of eq. (2), without further considerations.

### 1.2 Informal derivation of Kac-Rice

Consider a smooth compact manifold  $\mathcal{M}$  of dimension  $n$  (think of the unit sphere in  $n$  dimensions  $\mathbb{S}^{n-1}$ ), equipped with a Riemannian metric, and an associated volume measure  $\mu_{\mathcal{M}}$ . We are given a random smooth function  $f : \mathcal{M} \rightarrow \mathbb{R}$ . We want to use the area formula (Proposition 1.1) to estimate the moments of the number of critical points of  $f$ . Given the hypotheses of Proposition 1.1, a reasonable hypothesis is to assume that  $f$  is almost surely a Morse function, i.e. that *all its critical points are non-degenerate*. Since  $\mathcal{M}$  is compact, one easily deduces that the number of critical points of a Morse function is finite<sup>1</sup>. For any  $k \in \mathbb{N}$

---

<sup>1</sup>Note that the numbers of critical points of different indices of a Morse function are constrained by the topology of  $\mathcal{M}$  by the Morse inequalities, see [MSWW63] for a review on Morse theory.

and Borel set  $B \subseteq \mathbb{R}$ , we define  $\text{Crit}_{f,k}(B)$  to be the number of critical points  $x \in \mathcal{M}$  of  $f$  such that  $f(x) \in B$  and such that the index of  $\text{Hess } f(x)$  (that is the number of strictly negative eigenvalues of the Hessian) is equal to  $k$ . The informal area formula of eq. (2), applied to  $\text{grad } f$ , would read:

$$\text{Crit}_{f,k}(B) = \int_{\mathcal{M}} d\mu_{\mathcal{M}}(x) \delta(\text{grad } f(x)) |\det \text{Hess } f(x)| \mathbb{1}[f(x) \in B, i(\text{Hess } f(x)) = k] \quad (4)$$

Taking the expectation of this equality, one directly obtains the Kac-Rice formula:

**Proposition 1.2 (Kac-Rice formula, informal)** *Let  $\mathcal{M}$  be a smooth compact Riemannian manifold of dimension  $n$ , with volume measure  $\mu_{\mathcal{M}}$ . Let  $f : \mathcal{M} \rightarrow \mathbb{R}$  be a random function that is almost surely Morse. Denote  $\varphi_{\text{grad } f(x)}(0)$  the density of  $\text{grad } f(x)$  with respect to the Lebesgue measure on  $\mathbb{R}^{n-1}$ , taken at 0. Then:*

$$\mathbb{E} \text{Crit}_{f,k}(B) = \int_{\mathcal{M}} d\mu_{\mathcal{M}}(x) \mathbb{E} [|\det \text{Hess } f(x)| \mathbb{1}[f(x) \in B, i(\text{Hess } f(x)) = k] | \text{grad } f(x) = 0] \varphi_{\text{grad } f(x)}(0).$$

**Some remarks :**

1. The rigorous derivation of this formula is much more involved, as one has to start from the weak equality of Proposition 1.1 and to use continuity arguments in order to obtain an equality at  $u = 0$ , see for instance the proof of Kac-Rice performed in [AW09]. This is the first reason for which the formula is usually stated for Gaussian random fields, since in these cases these continuity arguments can be justified using simple hypotheses on the covariance of the random field. The second reason is that in general, conditional expectations of non-Gaussian random variables are intractable, making the Kac-Rice formula effectively useless since one has to know the law of the Hessian conditioned by the gradient being zero. Under many heavy technical conditions, one can however derive rigorous non-Gaussian versions of the Kac-Rice formula, see for example Theorem 12.1.1 of [AT09] and Theorem 6.7 of [AW09].
2. The Kac-Rice formula transforms a random differential geometry problem into a random matrix theory problem. This does not necessarily mean that the resulting problem will be simpler ! The main difficulty in evaluating the Kac-Rice formula comes from the distribution of the Hessian conditioned by the gradient being zero.
3. The Kac-Rice formula can be generalized to compute higher moments of the variable  $\text{Crit}_{f,k}(B)$  as well (see Theorem 6.3 of [AW09]), and can therefore be used to compute the second moment of the complexity (see [S<sup>+</sup>17] for an application to the pure spherical  $p$ -spin model) as well as perform heuristic replica calculations to obtain the quenched complexity (see for instance [RABC19]). Via Morse's theory, it can also be used to compute the moments of the Euler characteristic of the level sets of  $f$ , see [AA<sup>+</sup>13] for an example.

## 2 The annealed complexity of the pure spherical $p$ -spin model

We essentially detail here a calculation performed first by physicists, and made rigorous in [AAČ13]. We follow here the derivation of this last paper. For (anterior) theoretical physics derivation of the complexity of similar models using the Kac-Rice formula along with heuristic theoretical physics arguments (giving nevertheless the exact result), one can for instance read the works of Fyodorov: [Fyo04], [FW07].

### 2.1 Statement of the problem

Consider  $N \geq 1$ ,  $p \geq 3$ , and consider the following function  $f_{N,p}$  on the unit sphere  $\mathbb{S}^{N-1}$ :

$$f_{N,p}(\boldsymbol{\sigma}) \triangleq \sum_{1 \leq i_1, \dots, i_p \leq N} J_{i_1, \dots, i_p} \sigma_{i_1} \cdots \sigma_{i_p} \quad (\boldsymbol{\sigma} \in \mathbb{S}^{N-1}), \quad (5)$$

in which  $J_{i_1, \dots, i_p} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ . In the physics language, the Hamiltonian  $H_{N,p}$  of the spherical  $p$ -spin model (defined on  $\mathbb{S}^{N-1}(\sqrt{N})$ ), is related to  $f_{N,p}$  by:  $f_{N,p}(\boldsymbol{\sigma}) = \frac{1}{\sqrt{N}} H_{N,p}(\sqrt{N}\boldsymbol{\sigma})$ . For any Borel set  $B \subseteq \mathbb{R}$ , we want to compute the large  $N$  limit of the expectation of the number of local minima  $\boldsymbol{\sigma}$  of  $f_{N,p}$ , such that  $f_{N,p}(\boldsymbol{\sigma}) \in \sqrt{N}B$ , that we denote  $\text{Crit}_{N,p}^0(B)$ . One can apply the Kac-Rice formula (Proposition 1.2)<sup>2</sup>:

$$\mathbb{E} \text{Crit}_{N,p}^0(B) = \int_{\mathbb{S}^{N-1}} d\mu(\boldsymbol{\sigma}) \varphi_{\text{grad } f_{N,p}(\boldsymbol{\sigma})}(0) \mathbb{E} \left[ |\det \text{Hess } f_{N,p}(\boldsymbol{\sigma})| \mathbf{1} \left[ f_{N,p}(\boldsymbol{\sigma}) \in \sqrt{N}B, \text{Hess } f_{N,p}(\boldsymbol{\sigma}) \geq 0 \right] \middle| \text{grad } f_{N,p}(\boldsymbol{\sigma}) = 0 \right],$$

in which  $\mu$  is the usual surface measure on  $\mathbb{S}^{N-1}$ . Note that here grad and Hess denote the *Riemannian* gradient and Hessian on the sphere, while we will denote  $\nabla, \nabla^2$  the Euclidian gradient and Hessian.

## 2.2 The joint distribution of $(f(\boldsymbol{\sigma}), \text{grad } f(\boldsymbol{\sigma}), \text{Hess } f(\boldsymbol{\sigma}))$

Deriving the joint law of  $(f(\boldsymbol{\sigma}), \text{grad } f(\boldsymbol{\sigma}), \text{Hess } f(\boldsymbol{\sigma}))$  is a necessary step in the Kac-Rice method, as these three random variables appear in the conditional expectation. We fix  $\boldsymbol{\sigma} \in \mathbb{S}^{N-1}$ . In our case, it is immediate that the random variable  $(f(\boldsymbol{\sigma}), \text{grad } f(\boldsymbol{\sigma}), \text{Hess } f(\boldsymbol{\sigma}))$  is a Gaussian centered random variable, as the sum of independent Gaussian centered random variables. We thus simply need to compute its correlations to fully characterize their joint distribution. We naturally identify the tangent space  $\mathcal{T}_{\boldsymbol{\sigma}}(\mathbb{S}^{N-1})$  with  $\mathbb{R}^{N-1}$ . If we denote  $P_{\boldsymbol{\sigma}}^{\perp}$  the orthogonal projector on  $\{\boldsymbol{\sigma}\}^{\perp}$ , one has:

$$\text{grad } f(\boldsymbol{\sigma}) = P_{\boldsymbol{\sigma}}^{\perp} \nabla f(\boldsymbol{\sigma}), \quad (6)$$

$$\text{Hess } f(\boldsymbol{\sigma}) = P_{\boldsymbol{\sigma}}^{\perp} \nabla^2 f(\boldsymbol{\sigma}) P_{\boldsymbol{\sigma}}^{\perp} - \langle \boldsymbol{\sigma}, \nabla f(\boldsymbol{\sigma}) \rangle P_{\boldsymbol{\sigma}}^{\perp}. \quad (7)$$

For instance one can compute the covariance of the gradient:

$$\begin{aligned} \mathbb{E} [\text{grad } f(\boldsymbol{\sigma}) \text{grad } f(\boldsymbol{\sigma})^{\top}] &= P_{\boldsymbol{\sigma}}^{\perp} \mathbb{E} [\nabla f(\boldsymbol{\sigma}) \nabla f(\boldsymbol{\sigma})^{\top}] P_{\boldsymbol{\sigma}}^{\perp}, \\ &= p P_{\boldsymbol{\sigma}}^{\perp}. \end{aligned}$$

Using the same kind of simple algebraic calculations, one obtains the joint distribution:

**Lemma 2.1** *The joint law of  $(f(\boldsymbol{\sigma}), \text{grad } f(\boldsymbol{\sigma}), \text{Hess } f(\boldsymbol{\sigma}))$  is the following:*

$$\begin{cases} f(\boldsymbol{\sigma}) & \stackrel{d}{=} Z, \\ \text{grad } f(\boldsymbol{\sigma}) & \stackrel{d}{=} \sqrt{p} \mathbf{g}, \\ \text{Hess } f(\boldsymbol{\sigma}) & \stackrel{d}{=} \sqrt{2(N-1)p(p-1)} M_{N-1} - pZ \text{Id}_{N-1}, \end{cases} \quad (8)$$

in which  $Z \sim \mathcal{N}(0, 1)$ ,  $\mathbf{g} \sim \mathcal{N}(0, \text{Id}_{N-1})$ , and  $M_{N-1}$  is a matrix from the Gaussian Orthogonal Ensemble of size  $(N-1)$  with the convention  $\mathbb{E} M_{ij}^2 = \frac{1+\delta_{ij}}{2(N-1)}$ . The variables  $(Z, \mathbf{g}, M_{N-1})$  are pairwise independent.

Let us make the following important remarks:

1. The joint distribution of these variables is independent of  $\boldsymbol{\sigma}$ .
2. The variables  $(f, \text{Hess } f)$  are independent from  $\text{grad } f$ , so performing the conditioning in the Kac-Rice formula will be trivial.
3. From the gradient distribution, one easily obtains its density evaluated in 0:

$$\varphi_{\text{grad } f_{N,p}(\boldsymbol{\sigma})}(0) = e^{-\frac{N-1}{2} \ln(2\pi p)}.$$

---

<sup>2</sup>The proof that  $f_{N,p}$  is almost surely a Morse function can be found in [AAČ13]

Recalling that the volume of the unit sphere is  $V(\mathbb{S}^{N-1}) = 2\pi^{N/2}/\Gamma(N/2)$ , one deduces from the Kac-Rice formula and what we said above:

$$\mathbb{E} \text{Crit}_{N,p}^0(B) = \frac{2\pi^{N/2}}{\Gamma(N/2)} e^{-\frac{N-1}{2} \ln(2\pi p)} [2(N-1)p(p-1)]^{\frac{N-1}{2}} \mathbb{E} \left[ |\det H_{N-1}| \mathbf{1}(H_{N-1} \geq 0, z \in \sqrt{N}B) \right], \quad (9)$$

in which  $H_{N-1} \triangleq M_{N-1} - \sqrt{\frac{p}{2(N-1)(p-1)}} z$ , with  $z$  a standard Gaussian variable and  $M_{N-1}$  a GOE matrix of size  $N-1$  (see Lemma 2.1), independent of  $z$ . It is now completely clear that we reduced a random differential geometry problem (counting the number of critical points of a random smooth function) to a random matrix theory problem.

### 2.3 Simplification of the problem

It is clear from eq. (9) that the following lemma (from [AAČ13]) will be useful:

**Lemma 2.2** *Let  $G \subseteq \mathbb{R}$  a Borel set,  $X \sim \mathcal{N}(0, t^2)$  (for a  $t > 0$ ) and  $M_{N-1} \sim \text{GOE}(N-1)$ . Then:*

$$\begin{aligned} & \mathbb{E} [ |\det(M_{N-1} - X \text{Id}_{N-1})| \mathbf{1}((M_{N-1} - X \text{Id}_{N-1}) \geq 0, X \in G) ] \\ &= \frac{\Gamma\left(\frac{N}{2}\right) (N-1)^{-\frac{N}{2}}}{\sqrt{\pi t^2}} \mathbb{E}_{\text{GOE}(N)} \left[ e^{-\frac{N}{2} \left(\frac{1}{(N-1)t^2} - 1\right) \lambda_0^2} \mathbf{1}\left(\lambda_0 \in \sqrt{\frac{N-1}{N}} G\right) \right]. \end{aligned} \quad (10)$$

In this equation,  $\lambda_0$  is the smallest eigenvalue of a random matrix from the  $\text{GOE}(N)$  ensemble.

This lemma is a pure result of random matrix theory. We do not prove it here, as it is proven as a particular case of Lemma 3.3 of [AAČ13]. The idea of the proof is to write explicitly the joint law of the eigenvalues of  $M_{N-1}$ , denoted  $\lambda_1 \leq \dots \leq \lambda_{N-1}$ , and to interpret  $X$  as the smallest eigenvalue of a larger  $\text{GOE}(N)$  matrix with eigenvalues  $X \leq \lambda_1 \leq \dots \leq \lambda_{N-1}$ , and use Selberg's integral. Applying this lemma to eq. (9) with  $t = \sqrt{\frac{p}{2(N-1)(p-1)}}$  and  $G = (t\sqrt{N})B$  yields that for any  $N$ :

$$\mathbb{E} \text{Crit}_{N,p}^0(B) = 2\sqrt{\frac{2}{p}} (p-1)^{\frac{N}{2}} \mathbb{E}_{\text{GOE}(N)} \left[ e^{-N \frac{p-2}{2p} \lambda_0^2} \mathbf{1}\left(\lambda_0 \in \sqrt{\frac{p}{2(p-1)}} B\right) \right]. \quad (11)$$

### 2.4 The large $N$ limit

We are interested in the limit of the annealed complexity  $\Sigma_p^0(B) \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{E} \text{Crit}_{N,p}^0(B)$ . A precise look at eq. (11) can convince that we need to study the large deviations for the smallest eigenvalue of a  $\text{GOE}(N)$  matrix. We state this result in an informal way, see Theorem A.1 of [AAČ13] for a more precise statement, along with the Large Deviation Principle (LDP) for eigenvalues of all finite index:

**Lemma 2.3 (LDP of the smallest eigenvalue of a GOE matrix, informal)** *Let  $\lambda_0$  be the smallest eigenvalue of a  $\text{GOE}(N)$  matrix, and  $I \subseteq \mathbb{R}$ . Then*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P}[\lambda_0 \in I] = - \sup_{x \in I} F(x), \quad (12)$$

with  $F$  defined as:

$$F(x) \triangleq \begin{cases} \infty & \text{if } x \geq -\sqrt{2} \\ \frac{1}{\sqrt{2}} \int_{\sqrt{2}}^{-x} dz \sqrt{\frac{z^2}{2} - 1} & \text{otherwise} \end{cases}. \quad (13)$$

Using Lemma 2.3 alongside eq (11) and Varadhan's lemma directly yields a final estimate for the annealed complexity:

$$\Sigma_p^0(B) = \frac{1}{2} \ln(p-1) + \sup_{x \in \sqrt{\frac{p}{2(p-1)}} B} \left[ -\frac{p-2}{2p} x^2 - F(x) \right]. \quad (14)$$

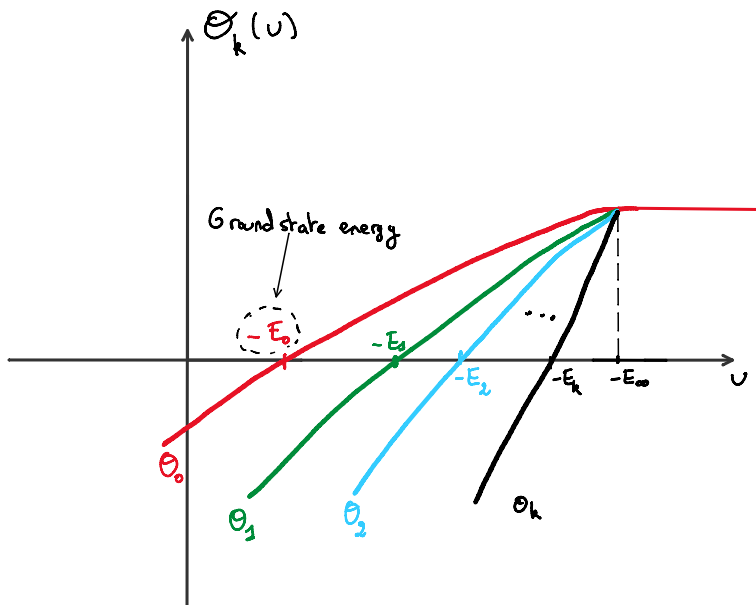


Figure 1: The functions  $\Theta_k$  for the first indices

## 2.5 Description of the results

Considering  $B = (-\infty, u)$  (for  $u \in \mathbb{R}$ ) in eq. (14) leads to effectively count the local minima of the pure  $p$ -spin Hamiltonian of extensive energy smaller than  $Nu$ . In this case, the supremum in eq. (14) can be analytically performed, and one obtains an explicit form for  $\Sigma_p^0((-\infty, u))$ . The calculation we sketched can also be performed for critical points of any fixed index  $k \in \mathbb{N}$  (see again [AAČ13]), and if we define:

$$\Theta_k(u) \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{E} \text{Crit}_{N,p}^k((-\infty, u)), \quad (15)$$

there exists analytic expressions for all  $\Theta_k(u)$  functions, see eq (2.16) of [AAČ13]. Defining the *threshold* energy  $E_\infty \equiv 2\sqrt{\frac{p-1}{p}}$ , we can sketch the plot of these functions, see Fig. 1. We also denote  $-E_k$  the value at which the function  $\Theta_k(u)$  becomes positive.

### Final remarks

1. The local minima always dominate the complexity for all energies below  $-NE_\infty$ , while for all energies above  $-NE_\infty$ , the complexity is dominated by critical points of diverging index.
2. The value  $-E_0$  is the lowest energy of critical points, and corresponds to the global minimum of the original function. It can be shown indeed ([AAČ13]) that the ground state energy concentrates to a deterministic value as  $N \rightarrow \infty$ .
3. One can perform similar calculations for the complexity of critical points whose indices diverge with  $N$ , see [AA+13] for the rigorous derivation.
4. This rigorous calculation can also be generalized to the mixed spherical  $p$ -spin case, see again [AA+13]. By a classical result of Schoenberg, these models exhaust all stationary isotropic Gaussian random fields on the sphere in  $N$  dimensions.

## References

- [AA<sup>+</sup>13] Antonio Auffinger, Gerard Ben Arous, et al. Complexity of random smooth functions on the high-dimensional sphere. *The Annals of Probability*, 41(6):4214–4247, 2013.
- [AAČ13] Antonio Auffinger, Gérard Ben Arous, and Jiří Černý. Random matrices and complexity of spin glasses. *Communications on Pure and Applied Mathematics*, 66(2):165–201, 2013.
- [AT09] Robert J Adler and Jonathan E Taylor. *Random fields and geometry*. Springer Science & Business Media, 2009.
- [AW09] Jean-Marc Azaïs and Mario Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, 2009.
- [Fed59] Herbert Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–491, 1959.
- [FW07] Yan V Fyodorov and Ian Williams. Replica symmetry breaking condition exposed by random matrix calculation of landscape complexity. *Journal of Statistical Physics*, 129(5-6):1081–1116, 2007.
- [Fyo04] Yan V Fyodorov. Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices. *Physical review letters*, 92(24):240601, 2004.
- [MSWW63] John Willard Milnor, Michael Spivak, Robert Wells, and Robert Wells. *Morse theory*. Princeton university press, 1963.
- [RABC19] Valentina Ros, Gerard Ben Arous, Giulio Biroli, and Chiara Cammarota. Complex energy landscapes in spiked-tensor and simple glassy models: Ruggedness, arrangements of local minima, and phase transitions. *Physical Review X*, 9(1):011003, 2019.
- [S<sup>+</sup>17] Eliran Subag et al. The complexity of spherical  $p$ -spin models—a second moment approach. *The Annals of Probability*, 45(5):3385–3450, 2017.