



HAL
open science

Mean-field theory for optimization problems

Henri Orland

► **To cite this version:**

Henri Orland. Mean-field theory for optimization problems. *Journal de Physique Lettres*, 1985, 46 (17), pp.763-770. 10.1051/jphyslet:019850046017076300 . jpa-00232896

HAL Id: jpa-00232896

<https://hal.science/jpa-00232896>

Submitted on 4 Feb 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LE JOURNAL DE PHYSIQUE - LETTRES

J. Physique Lett. 46 (1985) L-763 - L-770

1^{er} SEPTEMBRE 1985, PAGE L-763

Classification

Physics Abstracts

05.20 — 75.50 — 85.40

Mean-field theory for optimization problems

H. Orland

Service de Physique Théorique, CEN Saclay, 91191 Gif-sur-Yvette Cedex, France

(Reçu le 6 mai 1985, accepté le 8 juillet 1985)

Résumé. — On présente une théorie de champ moyen pour les problèmes d'optimisation du type « Voyageur de Commerce » ou problèmes d'« Appariement ». Cette théorie de champ moyen s'exprime à l'aide d'un ensemble infini de paramètres d'ordre, qui mesurent l'absence d'auto-moyennage du système et donc son degré de gel. On conjecture que la NP-complexité est associée à la brisure de symétrie des répliques.

Abstract. — A mean-field theory for optimization problems of the Travelling Salesman type, or of the Matching type, is presented. It involves an infinite set of order parameters which measure the lack of self-averageness of the system and its degree of freezing. We further conjecture that NP-completeness is associated with replica symmetry breaking.

1. Introduction.

Optimization problems [1] can often be expressed as the search for the minimum of a function of several variables $f(x_1, \dots, x_N)$, subject to constraints of the type $g_i(x_1, \dots, x_N) > 0$, $i = 1, \dots, m$.

These problems can be classified according to the time that the most efficient algorithms require to find the optimal solution. In the P-(polynomial) class, there exist algorithms for which the time grows polynomially with the « size » of the problem, whereas, in the NP-(non polynomial) class, the best known algorithms require a time which grows faster than any power of the size of the problem (typically exponential or factorial).

To be specific, in the following, we shall make use of two examples. In the P-class, we consider the Bipartite Matching Problem [1] (BMP) : given two sets of N points $\{ \mathbf{r}_i \}$ and $\{ \mathbf{p}_i \}$ in a d -dimensional space, and a distance l , find the best matching between the two sets, i.e. find the permutation

P of $\{ 1, 2, \dots, N \}$ such that $L = \sum_{i=1}^N l(\mathbf{r}_i, \mathbf{p}_{P(i)})$ be minimal. The problem can also be formulated as a cost problem in terms of a set of N^2 bonds l_{ij} and minimizing $L = \sum_{i=1}^N l_{iP(i)}$. These problems

can be solved in $\mathcal{O}(N^3)$ steps [1]. The BMP has been introduced as a physical model for line defects in crystals by Bachas [2]. In the NP-class, we consider the celebrated Travelling Salesman Problem (TSP) [3] : given N cities, find the shortest path which visits each city once and only once. It can be formulated either as a Euclidean problem, with N cities at points $\{\mathbf{r}_i\}$, with a distance l , and the quantity to minimize is the length $L = \sum_{i=1}^N l(\mathbf{r}_{P(i)}, \mathbf{r}_{P(i+1)})$, or as a cost problem as for the BMP.

The great difficulty in many optimization problems comes from the existence of a large number of local minima. This usually causes the algorithms to get trapped in a local minimum, with no certitude of reaching the absolute minimum.

The same difficulty exists in many physical systems (i.e. metastable states), particularly in Spin Glasses [4], where the number of metastable states is exponentially large [5]. To circumvent this difficulty, S. Kirkpatrick *et al.* [6] have proposed to use the methods of Statistical Mechanics, i.e. the Metropolis Monte-Carlo algorithm [7]. This procedure, called « simulated annealing » has been widely used by now [8], and turns out to be extremely efficient.

However, from an analytical point of view, very little has been done, except for the work of Vannimenus and Mézard [9] which have shown the existence of two regimes in the TSP and have pointed out some analogy with the spin-glass problem.

The purpose of this paper is to show clearly the analogy of some optimization problems with Spin Glass problems, and show how a mean-field theory can be derived.

2. The partition function.

In the following of the paper, we shall consider only cost (i.e. bond) problems. The partition functions of the BMP and TSP can be written as :

$$Z_{\text{BMP}} = \sum_{P \in S_N} \exp\left(-\beta \sum_{i=1}^N l_{iP(i)}\right) \quad (1)$$

and

$$Z_{\text{TSP}} = \frac{1}{2N} \sum_{P \in S_N} \exp\left(-\beta \sum_{i=1}^N l_{P(i)P(i+1)}\right) \quad (2)$$

where $P \in S_N$ denotes any permutation of $\{1, 2, \dots, N\}$, and $N+1 \equiv 1$.

Let us note :

$$U_{ij} = e^{-\beta l_{ij}}.$$

It is easily seen that :

$$Z_{\text{BMP}} = \int \prod_{i=1}^N d\eta_i^* d\eta_i d\xi_i^* d\xi_i \exp\left(-\sum_{i=1}^N \sum_{j=1}^N \eta_i^* \eta_i U_{ij} \xi_j^* \xi_j\right) \quad (3)$$

where the integral runs over Grassman (anticommuting) variables [10]. Indeed, using the basic rules for Grassman integrals [10] :

$$\int d\eta \cdot 1 = 0$$

$$\int d\eta \cdot \eta = 1$$

one sees on (3) that Z_{BMP} is a sum of all permutations where each site i is occupied exactly once

by the η_i^* η_i variables and by the ζ_i^* ζ_i variables, and there is a weight $U_{ij} = e^{-\beta l_{ij}}$ for each pair (i, j) .

Similarly, for the TSP, we want to construct closed polygons, and we are led to introduce p -vector spins S_i as in the High-Temperature expansion for spin systems :

$$\zeta = \int \prod_{i=1}^N d\mu(S_i) \prod_{\langle i,j \rangle} (1 + KU_{ij} S_i \cdot S_j) \tag{4}$$

where

$$d\mu(S_i) = dS_i \cdot \delta(S_i^2 - p).$$

Expanding ζ in powers of K , we see that it consists of a sum of polygons, with a factor KU_{ij} per bond, and a factor p per connected part. If we want the polygons to go through the N points, we must select the terms proportional to K^N , and if we want to keep only connected polygons, we must take the limit $p \rightarrow 0$, as in polymer theory [11, 12]. We have thus :

$$Z_{\text{TSP}} = \lim_{K \rightarrow +\infty} \lim_{p \rightarrow 0} \frac{1}{K^N} \cdot \frac{1}{p} \cdot \zeta.$$

The extraction of the K^N term can also be made by treating K as a chemical potential, and requiring that :

$$N = K \frac{\partial}{\partial K} \left(\lim_{p \rightarrow 0} \frac{1}{p} \zeta \right).$$

Finally, using the results of reference [12], the $p = 0$ limit allows us to exponentiate the spin variables in (4) :

$$\zeta = \int \prod_{i=1}^N d\mu(S_i) \exp \left(K \sum_{\langle i,j \rangle} U_{ij} S_i \cdot S_j \right). \tag{5}$$

3. Replicas.

As usually in the study of disordered systems, we wish to average the free energy (i.e. $\log Z$) over the distribution of random variables.

Following the techniques used in the theory of spin-glasses [4], we introduce replicas.

We assume that the U_{ij} are independent random variables, distributed with the same probability distribution $P(U_{ij})$, and we define the cumulants of the distribution P by :

$$\int dUP(U) e^{xU} = \exp \left(\sum_{k=1}^{\infty} \varphi_k \frac{x^k}{k!} \right). \tag{6}$$

The average of (3) and (5) yields :

$$\begin{aligned} \overline{Z_{\text{BMP}}^n} &= \int \prod_{\alpha=1}^n \prod_{i=1}^N d\eta_i^{*\alpha} d\eta_i^\alpha d\zeta_i^{*\alpha} d\zeta_i^\alpha \times \\ &\times \exp \left(\sum_{k=1}^{\infty} \varphi_k \sum_{1 \leq \alpha_1 \leq \dots \leq \alpha_k \leq n} \left(\sum_{i=1}^N |\eta_i^{\alpha_1}|^2 \dots |\eta_i^{\alpha_k}|^2 \right) \left(\sum_{j=1}^N |\zeta_j^{\alpha_1}|^2 \dots |\zeta_j^{\alpha_k}|^2 \right) \right) \end{aligned} \tag{7}$$

and

$$\overline{\zeta^n} = \int \prod_{\alpha=1}^n \prod_{i=1}^N d\mu(S_i^\alpha) \times \exp \left(\sum_{k=1}^{\infty} \frac{\varphi_k}{2 \cdot k!} \sum_{\alpha_1, \dots, \alpha_k=1}^n \sum_{\gamma_1, \dots, \gamma_k=1}^n \left(\sum_{i=1}^N S_{i\gamma_1}^{\alpha_1} \dots S_{i\gamma_k}^{\alpha_k} \right)^2 \right). \tag{8}$$

Let us note that in (7), the terms where two α 's coincide vanish, due to the anticommutation of the η 's and ξ 's variable. From now on, we shall focus on the BMP. Using a Gaussian transform, we get :

$$\overline{Z_{\text{BMP}}^n} = \int \prod_{k=1}^{\infty} \prod_{1 \leq \alpha_1 < \dots < \alpha_k \leq n} d\mu(q^*, q) \exp\left(-\sum_{k=1}^{\infty} \sum_{\alpha_1 < \dots < \alpha_k} \frac{|q_{\alpha_1 \dots \alpha_k}|^2}{\varphi_k}\right) \times e^{N \text{Log}(I(q^*) I(q))} \quad (9)$$

where

$$I(q) = \int \prod_{\alpha=1}^n d\eta_{\alpha}^* d\eta_{\alpha} \exp\left(\sum_{k=1}^{\infty} \sum_{\alpha_1 < \dots < \alpha_k} q_{\alpha_1 \dots \alpha_k} |\eta_{\alpha_1}|^2 \dots |\eta_{\alpha_k}|^2\right). \quad (9a)$$

4. Mean-field theory.

Mean field is obtained by evaluating (9) by the Saddle-Point method.

The Saddle-Point equations are :

$$q_{\alpha_1 \dots \alpha_k} = N\varphi_k \langle |\xi_{\alpha_1}|^2 \dots |\xi_{\alpha_k}|^2 \rangle \quad (10a)$$

$$q_{\alpha_1 \dots \alpha_k}^* = N\varphi_k \langle |\eta_{\alpha_1}|^2 \dots |\eta_{\alpha_k}|^2 \rangle \quad (10b)$$

To proceed further, we shall use recent results of the theory of spin-glasses [14]. In the infinite range model, we know that above the de Almeida-Thouless [15] line, in the (H, T) plane, the system is made of a unique phase, and there is no replica symmetry breaking, whereas below, there is an infinite (e^N) number of metastable phases, and there is replica symmetry breaking.

It seems plausible that an optimization problem is in the P-class if the number of local minima grows like a power of N , whereas it is in the NP-class if it grows faster than any power of N .

To proceed further, one must make an ansatz on the structure of the overlaps $q_{\alpha_1 \dots \alpha_k}$ in (10a). By analogy with spin-glasses, we assume :

i) No symmetry breaking for P-class problems.

ii) An ultrametric « à la Parisi » type of symmetry breaking for NP-problems [16].

In the BMP, we thus assume

$$q_{\alpha_1 \dots \alpha_k} = q_{\alpha_1 \dots \alpha_k}^* = q_k \quad \text{for } \alpha_1 < \dots < \alpha_k.$$

It is easily seen that $\left(\frac{q_k}{N\varphi_k}\right)^2 = \overline{\left(\frac{Z_2}{Z}\right)^k}$, where Z_2 is the partition function, given that two points are connected (the bar denotes as usual the average over the bond distribution). If the system were self-averaging, one would expect that $\left(\frac{q_k}{N\varphi_k}\right)^2 = \left(\frac{q_1}{N\varphi_1}\right)^k$, and $\left(\frac{q_1}{N\varphi_1}\right)^2$ would be independent of the particular instance of bonds.

A lack of self-averageness can be measured by :

$$\Delta_k = \left(\frac{q_k}{N\varphi_k}\right)^2 - \left(\frac{q_1}{N\varphi_1}\right)^{2k}.$$

Thus, a non-vanishing finite Δ_k signals the absence of self-averageness, i.e. the freezing of the system [14].

The analytic continuation at $n = 0$ is easily done, and the free energy functional reads :

$$-\beta F = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \frac{q_k^2}{\varphi_k} + 2N \int_0^{\infty} \frac{dt}{t} \left(e^{-t} - \exp\left(\sum_{k=1}^{\infty} (-1)^k q_k \frac{t^k}{k!} \right) \right). \quad (11)$$

The Saddle-Point method yields :

$$q_k = N \varphi_k \int_0^{\infty} dt \frac{t^{k-1}}{(k-1)!} \exp\left(\sum_{l \geq 1} (-1)^l q_l \frac{t^l}{l!} \right). \quad (12)$$

Denoting the generating function by :

$$\phi(x) = \sum_{k=1}^{\infty} (-1)^k q_k \frac{x^k}{k!} \quad (13)$$

the mean-field equations read :

$$\phi(x) = N \int_0^{\infty} \frac{dt}{t} f(xt) e^{\phi(t)} \quad (14)$$

with

$$f(x) = \sum_{k=1}^{\infty} (-1)^k \frac{\varphi_k}{k!} \cdot \frac{x^k}{(k-1)!} \quad (15)$$

or equivalently :

$$q_k = -N \frac{\varphi_k}{k!} \int_0^{\infty} t^k d(e^{\phi(t)}).$$

Let us note that the q_k are positive, and they satisfy the sum rule :

$$\sum_{k=1}^{\infty} (-1)^k \frac{q_k^2}{\varphi_k} = -N$$

and the inequalities :

$$k! \frac{q_k}{N \varphi_k} \geq \left(\frac{q_1}{N \cdot \varphi_1} \right)^k.$$

5. Discussion.

As in reference [9], we discuss the case of a Poisson distribution for the bonds l_{ij} :

$$P(l_{ij}) = \frac{l_{ij}^r}{r!} e^{-l_{ij}} \quad (16)$$

where $r = d - 1$, and d mimicks a dimension for the bond problem. As an illustration, we compute $\overline{Z^2}$. We note :

$$g_k(\beta) = \int_0^{\infty} dl P(l) e^{-k\beta l} = (1 + k\beta)^{-d}$$

and thus :

$$\begin{aligned}\varphi_1(\beta) &= g_1(\beta) \\ \varphi_2(\beta) &= g_2(\beta) - g_1^2(\beta).\end{aligned}\tag{17}$$

Using the formalism of the previous section, the average of $\overline{Z^2}$ is given by :

$$\overline{Z^2} = \text{Min}_{\{q_1, q_2\}} \left(\exp \left(-2 \frac{q_1^2}{\varphi_1} - \frac{q_2^2}{\varphi_2} + 2 N \text{Log} (q_1^2 + q_2) \right) \right)\tag{18}$$

and the mean field equations read :

$$q_1 = N\varphi_1 \cdot \frac{q_1}{q_1^2 + q_2}\tag{19a}$$

$$q_2 = N\varphi_2 \cdot \frac{1}{q_1^2 + q_2}.\tag{19b}$$

There is a critical point given by :

$$\frac{g(2\beta^*)}{g^2(\beta^*)} = N + 1\tag{20}$$

i.e.

$$\beta^* = 2 N^{1/d}, \quad T^* = \frac{1}{2} N^{-1/d}$$

i) if $T > T^*$,

$$\begin{aligned}q_2 &= \frac{\varphi_2(\beta)}{\varphi_1(\beta)} \\ q_1^2 &= N\varphi_1 - \frac{\varphi_2}{\varphi_1}\end{aligned}\tag{21}$$

$$\overline{Z^2} = [N ! g_1^N(\beta)]^2 \cdot \exp \left(2 \cdot \frac{\varphi_2(\beta)}{\varphi_1^2(\beta)} \right)$$

in agreement with reference [9], and the extensive part of $\overline{Z^2}$ is equal to the square of the annealed partition function :

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \text{Log} \left(\frac{\overline{Z^2}}{(\overline{Z})^2} \right) = 1$$

for any finite temperature T .

ii) if $T < T^*$,

$$\begin{aligned}q_2^2 &= N\varphi_2 \\ q_1 &= 0 \\ \overline{Z^2} &= N ! \varphi_2^N(\beta).\end{aligned}\tag{22}$$

The interpretation of this is quite simple : denoting by $\hat{T}^d = NT^d$, in the high temperature region, $T > T^*$, $\Delta_2 \simeq \frac{1}{2^d \hat{T}^{3d}} \left(2 - \frac{1}{2^d \hat{T}^d} \right)$ and there is self-averageness down to temperatures of order $N^{-\frac{1}{d}}$. In the low temperature region, $\Delta_2 = \left(\frac{2}{\hat{T}} \right)^d$, and there is a gradual freezing of the system.

If we study the case of \overline{Z}^n , $n \rightarrow 0$, the mean-field equations (12) read (in the region $\beta \gg 1$) :

$$q_k = \frac{NT^d}{k^r k!} \int_0^\infty dt t^{k-1} \exp\left(\sum_{l \geq 1} (-1)^l q_l \frac{t^l}{l!}\right). \tag{23}$$

We note

$$\hat{T}^d = NT^d. \tag{24}$$

Let us assume that one can neglect all the terms of order $k > 1$ in the exponent of (23), i.e. that the function $\phi(x)$ defined in equation (13) can be approximated by a linear function :

$$\phi(x) = -q_1 x. \tag{25}$$

This implies :

$$q_1^2 = \hat{T}^d \tag{26a}$$

$$q_k = \frac{\hat{T}^d}{(k\hat{T}^{k/2})^d} \tag{26b}$$

and

$$\Delta_k = 0. \tag{26c}$$

We see that as long as condition (25) holds, the system is self-averaging. However, when \hat{T} becomes smaller than 1, condition (25) cannot hold, since q_k increases when the temperature decreases, and that might be the sign for a phase transition, around $\hat{T} = 1$, or for a crossover at a higher temperature. In any case, as noted by Vannimenus and Mézard [9], there is a change of regime in the region $\hat{T} = 1$.

In the high temperature region, $\hat{T} \gg 1$, (i.e. T finite) we obtain :

$$-\beta F = -N + N \text{Log } q_1^2 = N \text{Log} \left(\frac{NT^d}{e} \right).$$

which is the annealed result.

In the low temperature region, the order parameters q_k are not small, and there is some freezing of the system.

We have not been able to find a mean-field solution in the low temperature region, and a numerical study is in preparation.

It might also be that in the low temperature region, the ansatz of no replica symmetry breaking might be wrong. Regarding the question of replica symmetry breaking, let us note that the most natural structure would be a Parisi type of structure for $q_{\alpha\beta}$, and the shape of $q_{\alpha_1 \dots \alpha_k}$ would be deduced by assigning it a value which depends on the clusters to which $\alpha_1 \dots \alpha_k$ belong. However, this introduces combinatorial difficulties which have not yet been solved.

Acknowledgments.

The author wishes to thank E. Brézin for useful discussions and C. De Dominicis for his encouragements and help.

References

- [1] PAPANIMITRIOU, C. H., STEIGLITZ, K., *Combinatorial Optimization* (Prentice Hall) 1982.
- [2] BACHAS, C. P., *Phys. Rev. Lett.* **54** (1985) 53.
- [3] BEARDWOOD, J., HALTON, J. H., HAMMERSLEY, J. M., *Proc. Camb. Philos. Soc.* **55** (1959) 299.
- [4] EDWARDS, S. F., ANDERSON, P. W., *J. Phys. F* **5** (1975) 965.
- [5] BRAY, A. J. and MOORE, M. A., *J. Phys. C* **13** (1980) L469.
DE DOMINICIS, C., GABEY, M., GAREL, T., ORLAND, H., *J. Physique* **41** (1980) 923.
TANAKA, F. and EDWARDS, S. F., *J. Phys. F* **10** (1980) 2769.
- [6] KIRKPATRICK, S., GELATT, C. D., Jr., VECCHI, M. P., *Science* **220** (1983) 671.
KIRKPATRICK, S., *J. Stat. Phys.* **74** (1984) 975.
- [7] BINDER, K., ed., *The Monte Carlo Method in Statistical Physics* (Springer, Berlin) 1978.
- [8] SIARRY, P., DREYFUS, M., *J. Physique Lett.* **45** (1984) L-139.
BONOMI, E., LUTTON, J. L., *SIAM Rev.* **26** (1984) 551.
KIRKPATRICK, S., TOULOUSE, G., to appear in *J. Physique*.
- [9] VANNIMENUS, J., MÉZARD, M., *J. Physique Lett.* **45** (1984) L-1145.
- [10] ITZYKSON, C., ZUBER, J.-B., *Quantum Field Theory* (McGraw-Hill) 1980, p. 439.
- [11] DE GENNES, P. G., *Phys. Lett. A* **38** (1972) 339.
- [12] DAOU, M., COTTON, J. P., FARNOUX, B., JANNINK, G., SARMA, G., BENOIT, H., DUPLESSIX, R.,
PICOT, C. and DE GENNES, P. G., *Macromolécules* **8** (1975) 804.
- [13] PARISI, G., *Phys. Rev. Lett.* **50** (1983) 1946.
- [14] MÉZARD, M., PARISI, G., SOURLAS, N., TOULOUSE, G., VIRASORO, M., *Phys. Rev. Lett.* **52** (1984) 1156
and *J. Physique* **45** (1984) 843.
- [15] DE ALMEIDA, J. R. L., THOULESS, D. J., *J. Phys. A* **11** (1978) 983.
- [16] PARISI, G., *Phys. Rev. Lett.* **43** (1979) 1754 and *J. Phys. A* **13** (1980) L115, 1101.
-