

# Article on statistical mechanics methods and phase transitions in optimization problems.

Christian Vestergaard

June 11, 2007

## 1 Introduction

Theoretical methods from statistical physics can be used in optimization problems in computer science and mathematics, as they often be defined by an energy-function which needs to be minimized to find the optimal solution. These methods can be used to gain information about statistical properties of these systems, and can aid in finding the lowest energy configuration (i.e. optimization). In addition statistical mechanics can help understand the phase-transitions which occur in a lot of optimization problems. Statistical mechanics has also given inspiration to a generic algorithm for finding the optimal or near-optimal solution to a complex optimization problem, simulated annealing.

## 2 Discussion

In the context of computer science, statistical physics methods are often used on the NP-hard problems, where the NP classification means that no optimization algorithm has been found that is polynomial in the max. time it used to resolve the problem. I.e. an algorithm classified as polynomial (P) uses in the worst case scenario a time that is polynomial in the size of the system  $N$ ,  $t = O(N^p)$ . For the NP problems, the worst case time is exponential in  $N$ . The article goes into detail with 3 different problems in optimization:

- Random graphs
- Random K-satisfiability
- The stochastic travelling salesman problem

Due to lack of space, I will only talk about the first and last.

### 2.1 Random graphs

The article starts out showing how statistical physics methods can be used to analyze random graphs, equivalent to percolation in the  $N \rightarrow \infty$  limit. The random graph is used as a model problem, by mapping to the so-called Pott's model, statistical physics results are obtained.

A random graph  $G$  can be defined by a set of  $N$  points, there are  $N(N-1)/2$  edges (possible connections), where each edge is occupied with the probability

$p = \gamma/N$ , so in total on average  $N_L = \gamma N/2$  edges are occupied. (As  $N$  is large  $N/2 = (N-1)/2$ ).  $G$  is the n a random realization of these  $\gamma N/2$  connections.

The probability of a given random graph is:

$$P(G) = p^{N_L(G)}(1-p)^{N(N-1)/2-N_L(G)}$$

We want to get information about the average number of clusters per point  $c(G) = C(G)/N$ , that is we want to analyze the probability distrution of  $c$ ,  $\rho$ . To facilitate a statistical mechanic analysis of  $\rho$ , a mapping to the Pott's model, an Ising like model, is used. The Pott's model is defined by the Hamiltonian

$$\mathcal{H}[\{\sigma_i\}_i] = - \sum_{i<j} \delta_{\sigma_i, \sigma_j}, \quad \sigma_i \in \{0, 1, \dots, q-1\}$$

To allow mapping onto the Pott's model a cluster generating function  $Y(q)$  is defined, and is shown in the thermodynamic limit ( $N \rightarrow \infty$ ) to be related to  $\rho(c)$  by a Legendre transform. This is shown using a saddle-point calculation, which is only valid at large  $N$ . That is, if we define

$$\omega(c) = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \rho(c), \quad \tilde{f}(q) = \lim_{N \rightarrow \infty} \frac{1}{N} \ln Y(q)$$

we get from the saddle-point calculation

$$\tilde{f}(q) = \max_{0 \leq c \leq 1} [c \ln q + \omega(c)]$$

By expanding and rewriting the partition function  $Z_{Pott's}$  for the Pott's model it can be shown to be equal to  $Y(q)$ , except for a multiplicative factor.

$$Z_{Pott's} = e^{\gamma N/2} Y(q).$$

Defining  $f_{Pott's}(q) = -\lim_{N \rightarrow \infty} \frac{1}{N} \log Z_{Pott's}$  leads to the following equation, linking  $f_{Pott's}$ ,  $q$ ,  $c$  and  $\omega$ :

$$\max_{0 \leq c \leq 1} [c \ln q + \omega(c)] = -\frac{\gamma}{2} - f_{Pott's}(q)$$

This has to be maximized with respect to  $q$ , furthermore the r.h.s. is to be maximized with respect to  $c$  to find the stable solution  $c^*$ :

$$\left. \frac{d\omega(c)}{dc} \right|_{c^*} = -\ln q \quad \Rightarrow \quad q = 1.$$

and

$$c^*(q) = -q \frac{df_{Pott's}}{dq}(q).$$

Since  $q$  is originally defined as an integer value, an analytical of  $q$  to real values is needed to evaluate this expression. This is done by rewriting  $Z_{Pott's}$  in terms of the fraction  $x(\sigma; \{\sigma_i\}_i)$  of the  $\sigma_i$ 's in each of the  $q$  different states. And in the large- $N$  limit  $f_{Pott's}$  is found to be

$$f_{Pott's}(q) = \min_{\{x(\sigma)\}_i} \sum_{\sigma=0}^{q-1} \left\{ -\frac{\gamma}{2} [x(\sigma)]^2 + x(\sigma) \ln x(\sigma) \right\} \equiv \min_{\{x(\sigma)\}_i} f[\{x(\sigma)\}]$$

$f$  is symmetric under permutation of the  $q$ -values, thus we look firstly at the symmetric extremum for  $f$ :  $x^{sym}(\sigma) = 1/q$ , where it is found that  $c^{sym}(\gamma) = 1 - \gamma/2$ . This saddle-point cannot exist for  $\gamma > 2$  though, computation of the Hessian matrix (the second derivative as  $q \rightarrow 1$ ) of  $f$  at this symmetric saddle-point gives the eigenvalue  $\lambda = 1 - \gamma$ , which changes sign at  $\gamma_c = 1$  where the saddle-point becomes unstable and a percolating cluster appears.

By proposing unsymmetric saddle-points for  $\gamma > 1$  is the type

$$\begin{aligned} x(0) &= \frac{1}{q}[1 - (1 - q)s] \\ x(\sigma) &= \frac{1}{q}[1 - s] \end{aligned}$$

and taking the  $q \rightarrow 0$  limit, an implicit equation for the stationary solution of  $s$  is obtained:

$$1 - s^* = \exp(-\gamma s^*)$$

## 2.2 The stochastic travelling salesman problem (stochastic TSC)

The travelling salesman problem is a very well known optimization problem and is NP-hard. A salesman need to go to  $N$  different cities, visiting each one exactly one time. The cities have randomly distributed individual distances, the problem is then to find the optimal route, i.e. the shortest total distance. The energy which needs to be minimized is very naturally defined as the length of the total distance.

There are different TSP's, the most well-known is the euclidian TSP, where the cities are independent randomly distributed points in  $d$ -dimensional euclidian space, the  $N(N - 1)/2$  distances between the cities are then given and obey the triangle inequality, it is thus called a 'metric' TSP. This problem is difficult from a statistical approach, since the distances are not independent. The random-distance TSP is chosen instead, where the distances between the cities are i.i.d. random variables, this problem is not metric, though still similar in many aspects to the euclidian TSP.

It has been shown that given a pretty well-behaved distribution  $\rho$  of the points in the eucliden TSP, the optimum tour length  $L_{min}$  is self-averaging and for  $N \rightarrow \infty$

$$\frac{L_{min}}{N^{1-1/d}} \rightarrow \beta(d) \int_{\Omega} \rho^{1-1/d}(\mathbf{X}) d\mathbf{X}$$

The problem is now to find  $\beta(d)$ .

For a stat. mechanics approach we need to define an energy function, a very obvious choice is the tour length  $L(\sigma)$ , where  $\sigma$  is the specific configuration of vertices that determines the length. This allows us to define

$$Z(T) = \sum_{\sigma} \exp\left(-\frac{L(\sigma)}{T}\right)$$

From this we can f.ex. extract the thermal average of the length  $\langle L \rangle = -\frac{d \ln Z}{d(1/T)}$ , and we see that  $L_{min} = \lim_{T \rightarrow 0} \langle L \rangle$ .

$Z$  is difficult to work with though, as it involves summing over all permutations of the distances, so a mapping to another spin-model is performed. An

m-dimensional vectorial spin  $\mathbf{S}$  is defined, with  $|\mathbf{S}|^2 = m$ . Each vertex of the graph has an associated spin  $\mathbf{S}_i$ , we define  $R_{ij} = e^{-d_{ij}/T}$  and introduce a new partition function

$$G(T, m, \omega) = \int \prod_{q=1}^N d\mathbf{S}_q \exp \left( \omega \sum_{i<j} R_{ij} \mathbf{S}_i \cdot \mathbf{S}_j \right)$$

It can be shown that

$$\lim_{\omega \rightarrow \infty} \left( \lim_{m \rightarrow 0} \frac{G - 1}{m\omega^N} \right) = Z(T)$$

This can be seen by Taylor-expanding the exponential in the equation for  $G$ :

$$G(T, m, \omega) = \int \prod_{q=1}^N d\mathbf{S}_q \left[ 1 + \omega \sum_{i<j} R_{ij} \mathbf{S}_i \cdot \mathbf{S}_j + \frac{1}{2} \omega^2 \dots \right]$$

and integrate over all  $\mathbf{S}_i$ , each contribution can be seen as a subgraph, but which doesn't have to be fully connected as in the TSP, also each vertex can have more than 2 two edges. Now due to that  $|\mathbf{S}_i|^2 = m$ , each loop will contribute a multiplicative factor  $m$ , so when  $m \rightarrow 0$  all the graphs with more than one loop will not contribute (graphs will have at least one loop). Then when  $\omega \rightarrow \infty$ , only the biggest loops will contribute, i.e. full tours. Furthermore the weight of each tour will be proportional to the product  $\prod_{i,j \in \sigma} \omega R_{ij}$ . so the total weight for each tour is  $m\omega^N \exp(-L/T)$ , and everything adds up.

This mapping allows us to define recursive equations for thermal expectation values.

By integrating  $G_{N+1} - 1$  over the  $N + 1^{\text{th}}$  spin we get a quantity depending only on the  $N$  first spins. By expanding the exponential we get a 1, which is eaten by the  $-1$ , a term linear linear in  $N + 1$ , which integrates to zero, thus the first non-zero term is the quadratic term, higher powers vanish as well when  $m \rightarrow 0$ . Thus we get:

$$\frac{Z_{N+1}}{Z_N} = \frac{G_{N+1} - 1}{G_N - 1} = \omega^2 \sum_{1 \leq j < k \leq N} R_{j,N+1} R_{k,N+1} \langle \mathbf{S}_j \cdot \mathbf{S}_k \rangle_N$$

In similar fashion recursive equations can be obtained for  $\langle \mathbf{S}_{N+1} \rangle_{N+1}$  and  $\langle n_{i,N+1} \rangle_{N+1} = \langle \mathbf{S}_{N+1} \cdot \mathbf{S}_i \rangle_{N+1}$ . Sadly, because of the correlation terms these equations cannot be solved, but by neglecting correlations between the first  $N$  spins the correlation terms factorise and the equations can be resolved.

As in the Ising model the spins are going to align spontaneously when the temperature becomes low enough, we can choose this direction by inducing a small electric field  $\mathbf{h}$  which we let go to zero after  $N \rightarrow \infty$ . If we denote  $S^1$  the component of  $\mathbf{S}$  in the direction of  $\mathbf{H}$ , we have the equations:

$$\begin{aligned} \langle S_{N+1}^1 \rangle_{N+1} &= \frac{\sum_j R_{j,N+1} \langle S_j^1 \rangle_N}{\omega \sum_{j<k} R_{j,N+1} R_{k,N+1} \langle S_j^1 \rangle_N \langle S_k^1 \rangle_N} \\ \langle n_{i,N+1}^1 \rangle_{N+1} &= \frac{R_{i,N+1} \langle S_i^1 \rangle_N \sum_{j \neq i} R_{j,N+1} \langle S_j^1 \rangle_N}{\sum_{j<k} R_{j,N+1} R_{k,N+1} \langle S_j^1 \rangle_N \langle S_k^1 \rangle_N} \end{aligned}$$

To have a thermodynamic limit though, we have to f.ex. rescale the temperature  $T = \tilde{T}N^{-1/d}$ . Now, by introducing a new function  $\phi_i = \tilde{T} \ln(\omega^{1/2} \langle S_i^1 \rangle_N)$  and a clever rearrangement of the variables we can obtain the distribution function for  $\langle S_i^1 \rangle_N$ .

It is very feasible that for the random-distance TSP the spins are in fact uncorrelated in the thermodynamic limit, and as thus the model is exact in this limit. Though it has not been proven numerical studies support this conjecture. For uniformly distributed distances the error between studies and prediction is at the level of 0.05%.

We can however not assume that the spins are uncorrelated in the Euclidean TSP, since neighbours neighbours are also neighbours. But in spite of this, the method actually gives a good approximation to the results in the Euclidean TSP. For  $d=2$  the cavity method predicts  $\beta(2) = 0.7251$ , while numerical simulations give  $0.7120 \pm 0.0004$ , which is really a small difference taking into account the approximations made.

In conclusion, the methods of statistical mechanics seem to work very well in a wide range of optimization problems, even though they are often not very well mathematically grounded.