A Statistical Mechanics Approach to Random Euclidean MAX TSP

David S. Dean(1), David Lancaster(2) and Satya.N. Majumdar(3)
(1) Laboratoire de Physique Théorique, UMR CNRS 5152, IRSAMC,
Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 04, France
(2) Harrow School of Computer Science, University of Westminster, Harrow, HA1 3TP, UK
(3) Laboratoire de Physique Théorique et Modèles Statistiques,
UMR 8626, Université Paris Sud, Bât 100, 91045 Orsay Cedex, France
(Dated: September 2005)

We present a statistical mechanics method to compute the average path length for random cases of the Euclidean TSP. This illuminates the problem from a new perspective, and in the case of MAX TSP reproduces some known results.

I. INTRODUCTION

The statistical mechanical approach to the study of combinatorial optimization problems has led to progress in a number of ways [1]. The approach is based on identifying the cost function, which needs to be minimized, with the energy of a physical system whose phase space is equivalent to the free adjustable parameters in the optimization problem. This physical system is then studied with the techniques of statistical mechanics at arbitrary temperature and the zero temperature configuration corresponds to the optimal solution. Here we are concerned with applying this statistical mechanical approach to the traveling salesman problem, in particular its Euclidean geometric version. The problem is to find the optimal (either minimal or maximal [2]) closed circuit length to visit N cities or points where the distance between the points i and j is given by d_{ij}. The order in which the cites are visited is encoded in a permutation $\sigma \in \Sigma_N$ where Σ_N, the group of permutations of N objects, acts as the phase space. For a given permutation

$$D(\sigma) = \sum_i d_{\sigma_i,\sigma_{i-1}}, \tag{1}$$

is the total distance traveled and corresponds to the cost function.

Our interest is in the stochastic case of the TSP in which the d_{ij}'s are chosen from some probability distribution. For the Euclidean case, this probability distribution is determined by taking the cities as points $\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N$ in some connected domain $\mathcal{D}$ in $\mathbb{R}^d$ where each point is independently distributed from the others with the same probability density function $p(\mathbf{r})$. The distance between the points i and j is simply the Euclidean distance on $\mathbb{R}^d$ given by $d_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. In this stochastic formulation, the quantities we are concerned with are averages over the locations of the cities. Our primary focus of attention is on computing the average optimal path length for randomly distributed cities in the limit of many cities. This quantity is of interest in its own right since path lengths for particular instances have been shown to be sharply peaked about this average [3], and it has been traditionally employed to test heuristics [4]. We also compute the number of paths of given length which characterizes some of the difficulty of finding the optimum one.

Statistical mechanical approaches to the TSP have been considered in the past, though most progress has been with the independent link version in which there is no correlation, such as the triangle inequality, between the random distances [5–9]. Recently we have devised a new statistical mechanical technique to compute the average path length for the Euclidean problem [10, 11]. Our method will be shown to be exact in the limit of large N while keeping the domain $\mathcal{D}$ fixed. The method yields integral equations that determine the average at any temperature, dimension and topology, but at present only for the leading term proportional to N, the number of cities, as $N \to \infty$. Since for the usual (minimal) TSP, this term vanishes at zero temperature, as the average minimal path length behaves as $N^{1-1/d}$ [3], our method is of most interest for the maximal TSP [2], where the average maximal path length does grow as N.

We first present our formalism and give some evidence that it is correct at finite temperature. We then proceed to investigate the zero temperature limit corresponding to the optimum and find that the integral equations reduce to minimization equations. A fascinating aspect of our approach is that the ansatz used to solve these zero temperature equations can be constructed by considering a simple class of algorithm for solving the TSP, notably the greedy algorithm.

II. FORMALISM

The fundamental object of the statistical mechanics approach is the partition function which acts as a kind of generating function. It is given by the following sum over all possible paths at temperature $1/\beta$ (for the maximal problem we take β to be negative).

$$Z_N = \frac{1}{N!}\sum_{\sigma\in\Sigma_N} \exp\left(-\beta D(\sigma)\right) = \frac{1}{N!}\sum_{\sigma\in\Sigma_N} \exp\left(-\beta\sum_i |\mathbf{r}^q_{\sigma_i} - \mathbf{r}^q_{\sigma_{i-1}}|\right). \tag{2}$$

Where the N points $\{\mathbf{r}^q_1, \mathbf{r}^q_2 \cdots, \mathbf{r}^q_N\}$ are chosen independently in some domain $\mathcal{D} \in \mathbb{R}^d$ with probability density $p(\mathbf{r})$ (the points have superscript q to distinguish them from the dummy integration variables introduced later). The path is made closed by defining $\mathbf{r}^q_0 = \mathbf{r}^q_N$. For this so-called "quenched" problem, it is incorrect to take the average over the location of the points at this stage, instead it should be performed after the average path length is computed. We take the leading term proportional to N and determine the average path length per city as:

$$E = -\frac{1}{N}\frac{\partial}{\partial\beta}\overline{\ln(Z_N)} \tag{3}$$

The average is denoted by an overline and rather than evaluating it explicitly it will be obtained by considering a large number of cities and requiring that their average density follows the desired $p(\mathbf{r})$.

To motivate our approach, consider a discrete problem consisting of M sites at points $\mathbf{r}^q_j$ (usually spaced regularly on a lattice), and imagine that there are n_j cities at site j. A path through the N cities is specified by the sequence of site indices $j_1, j_2 \dots j_N$, but to ensure that site j is visited precisely n_j times a Kronecker delta constraint is applied. Then for any function $f()$, we may write the sum over all permutations defining the paths as:

$$\sum_{\sigma\in\Sigma_N} f\left(\mathbf{r}^q_{\sigma_1}, \mathbf{r}^q_{\sigma_2}, \cdots \mathbf{r}^q_{\sigma_N}\right) = \sum_{j_1, j_2 \dots j_N=1}^{M} \prod_{k=1}^{M} n_k!\, \delta_{n_k, \sum_{m=1}^N \delta_{k,j_m}}\, f\left(\mathbf{r}_{j_1}, \mathbf{r}_{j_2}, \cdots \mathbf{r}_{j_N}\right) \tag{4}$$

We take a continuum version of this identity presuming that it is valid for sufficiently smooth functions.

$$\sum_{\sigma\in\Sigma_N} f\left(\mathbf{r}^q_{\sigma_1}, \mathbf{r}^q_{\sigma_2}, \cdots \mathbf{r}^q_{\sigma_N}\right) \propto \int \prod_{i=1}^{N} d^d r_i \prod_{\mathbf{r}} \delta\left(Np(\mathbf{r}) - \sum_i \delta(\mathbf{r}-\mathbf{r}_i)\right) f\left(\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_N\right) \tag{5}$$

Throughout, we take the domain size to be unity in order to simplify the equations. The occupancy factors, n_j, now appear as the specified density $p(\mathbf{r})$ of cities in space, so we have lost any explicit dependence on the locations, $\mathbf{r}^q_i$, of the original cities and in effect, have already performed the average.

We therefore write the partition function of the TSP as,

$$Z_N \propto \int \prod_{i=1}^{N} d^d r_i \prod_{\mathbf{r}} \delta\left(Np(\mathbf{r}) - \sum_i \delta(\mathbf{r}-\mathbf{r}_i)\right) \exp\left(-\beta\sum_i |\mathbf{r}_i - \mathbf{r}_{i-1}|\right). \tag{6}$$

The factor of proportionality is a temperature independent entropy/degeneracy contribution which can be determined by the discrete approach, or fixed by noting that $Z_N(0) = 1$ in the original definition.

Rewrite this using a Fourier representation of the functional constraint:

$$Z_N = \int d[\mu] \exp\left(N\int d^d r \mu(\mathbf{r}) p(\mathbf{r})\right) \mathcal{Z}_N, \tag{7}$$

where each integration over $\mu(x)$ is up the imaginary axis. The object $\mathcal{Z}_N$ defines a subsidiary statistical mechanical problem with partition function given by,

$$\mathcal{Z}_N = \int \prod_{i=1}^{N} d^d r_i \, \exp\left(-\beta\sum_{i=1}^{N} |\mathbf{r}_i - \mathbf{r}_{i-1}| - \sum_{i=0}^{N} \mu(\mathbf{r}_i)\right), \tag{8}$$

which will be solved below.

For large N we may evaluate the partition function in Eq. (7) by the saddle point method in the limit where $N \to \infty$ keeping $\mathcal{D}$ fixed. The saddle point equation is

$$p(\mathbf{r}) = -\frac{1}{N}\frac{\delta \ln \mathcal{Z}_N}{\delta \mu(\mathbf{r})} = \frac{1}{N}\langle \sum_{i=1}^{N} \delta(\mathbf{r}-\mathbf{r}_i)\rangle = p_a(\mathbf{r}) \tag{9}$$

where the above expectation (denoted by angled brackets), defining $p_a(\mathbf{r})$, is in the system with partition function $\mathcal{Z}_N$ given in Eq. (8). Physically this approach can be thought of as choosing a site dependent chemical potential μ which fixes the density of the (annealed) subsidiary problem to be the same as that of the specified (quenched) one $p(\mathbf{r}) = p_a(\mathbf{r})$.

To proceed, we find an expression for $p_a(\mathbf{r})$ by solving the subsidiary problem using transfer operator techniques.

$$\mathcal{Z}_N = \mathrm{Tr}\, \mathcal{T}^N \tag{10}$$

where

$$\mathcal{T}(\mathbf{r},\mathbf{r}') = \exp\left(-\mu(\mathbf{r})/2\right) \exp\left(-\beta|\mathbf{r}-\mathbf{r}'|\right) \exp\left(-\mu(\mathbf{r}')/2\right) \tag{11}$$

The ground state eigenfunction, corresponding to the maximal eigenvalue λ, obeys

$$f(\mathbf{r}) = \lambda^{-1} \exp\left(-\mu(\mathbf{r})/2\right) \int d^d r' \, \exp\left(-\beta|\mathbf{r}-\mathbf{r}'|\right) \exp\left(-\mu(\mathbf{r}')/2\right) f(\mathbf{r}') \tag{12}$$

and we have:

$$p_a(\mathbf{r}) = -\frac{\delta \ln(\lambda)}{\delta \mu(\mathbf{r})} = f^2(\mathbf{r}) \tag{13}$$

Returning to the saddle point equation, $p(\mathbf{r}) = p_a(\mathbf{r})$, we have that $f = \sqrt{p}$. This ensures that f is the eigenfunction corresponding to the maximal eigenvalue as we note that it is positive and then appeal to the Perron-Frobenius theorem.

Using the expression for $f(\mathbf{r})$ the saddle point equation becomes:

$$\sqrt{p(\mathbf{r})} = \lambda^{-1} \exp(-\frac{\mu(\mathbf{r})}{2}) \int d^d r' \, \exp\left(-\beta|\mathbf{r}-\mathbf{r}'|\right) \exp(-\frac{\mu(\mathbf{r}')}{2}) \sqrt{p(\mathbf{r}')} \tag{14}$$

we rewrite this equation in terms of $s_\lambda(\mathbf{r})$ given by $\exp(-\mu(\mathbf{r})/2) = \sqrt{p(\mathbf{r})}/s_\lambda(\mathbf{r})$.

$$s_\lambda(\mathbf{r}) = \lambda^{-1} \int d^d r' \, \exp\left(-\beta|\mathbf{r}-\mathbf{r}'|\right) \frac{p(\mathbf{r}')}{s_\lambda(\mathbf{r}')} \tag{15}$$

Using this in the value of the saddle point exponent we obtain,

$$\frac{\ln(Z_N)}{N} = 2\int d^d r \; p(\mathbf{r}) \ln\left(s_\lambda(\mathbf{r})\right) \; + \ln(\lambda) - \int d^d r \; p(\mathbf{r}) \ln(p(\mathbf{r})). \tag{16}$$

From Eq. (15) we see that there is a whole family of solutions, $\{s_\lambda(\mathbf{r}), \lambda\}$, related by $s_\lambda = a^{1/2} s_{a\lambda}$, for $a > 0$ and in addition these solutions all have the same action. This apparent zero mode is an artifact introduced by the fact that the constraint $N \int d^d r \; p(\mathbf{r}) = \int d^d r \; \sum_i \delta(\mathbf{r}-\mathbf{r}_i)$ is automatically satisfied. Thus we may chose $\lambda = 1$.

In the case of a uniform distribution on a domain of unit volume we have our final result for the average path length at inverse temperature β in terms of the quantity $s(\mathbf{r})$ determined by a non-linear integral equation.

$$\begin{aligned} E &= -2\frac{\partial}{\partial \beta}\left[\int d^d r \; \ln\left(s(\mathbf{r})\right)\right] \\ &= \int d^d r \; d^d r' \; \frac{|\mathbf{r}-\mathbf{r}'| \exp\left(-\beta|\mathbf{r}-\mathbf{r}'|\right)}{s(\mathbf{r}) s(\mathbf{r}')} \end{aligned} \tag{17}$$

where s obeys

$$s(\mathbf{r}) = \int d^d r' \; \frac{\exp\left(-\beta|\mathbf{r}-\mathbf{r}'|\right)}{s(\mathbf{r}')} \tag{18}$$

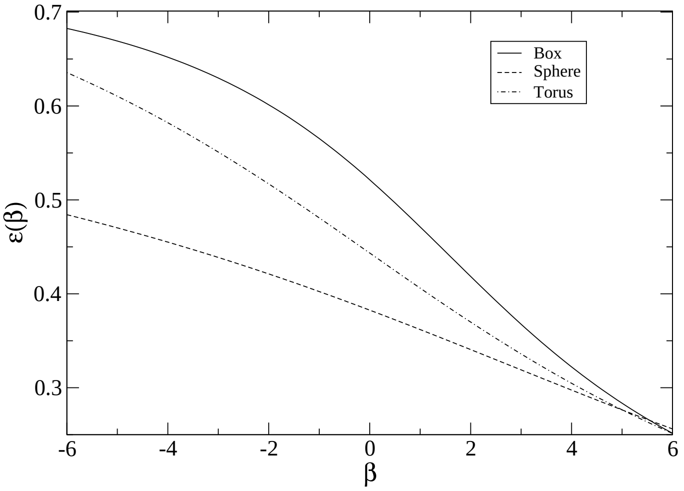

FIG. 1: Theoretical prediction for the average path length E for two dimensional TSP on a box, sphere and torus as a function of β. Negative β corresponds to the MAX TSP.

III. FINITE TEMPERATURE

The formalism presented above yields the correct high temperature expansion as a series in β, and has also been tested by comparison with Monte Carlo simulation of the TSP at finite temperature [11]. Except for the case of closed symmetric domains to be discussed below, we have not found any non-trivial analytic solutions of the nonlinear equations (17) and (18). For the special closed domains, such as a disc and torus in two dimensions, a constant solution exists. The significance of this observation is that the annealed approximation, in which the average over cities is taken at the level of the partition function Eq. (2), is exact for these domains.

For the sphere, the annealed/quenched equations have a constant solution with:

$$s^2 = \frac{1}{4\pi} \int_{S_2} d^2 r \ e^{-\beta\theta} = \frac{2\pi}{(\beta^2 + 4\pi)}(1 + e^{-\sqrt{\pi}\beta/2}) \tag{19}$$

leading to an expression for the average path length,

$$E = \frac{2\beta}{\beta^2 + 4\pi} + \frac{\sqrt{\pi}}{2(e^{\sqrt{\pi}\beta/2} + 1)} \tag{20}$$

Note that in the limit $\beta \to -\infty$, the path length becomes the half circumference corresponding, as is the case for all the closed domains, to the maximum distance two points can be apart.

For a torus we find that the equations yield

$$s^2 = \frac{8}{\beta^2} \int_0^{\pi/4} (1 - e^{-\beta/2\cos\theta} - \frac{\beta e^{-\beta/2\cos\theta}}{2\cos\theta})\, d\theta \tag{21}$$

As $\beta \to -\infty$, the integral can be approximated and again the path length becomes the maximum distance two points can be apart, $E \to 1/\sqrt{2}$.

For most domains, including the traditional square box domain, the quenched result is different from the annealed approximation and there is little hope of a general analytic solution. In this case, our primary tool is the iterative numerical solution of Eq. (18) which is stable and can be solved to any required accuracy.

Figure (1) shows the average path length for each of the three two-dimensional domains we have considered. For the sphere this is given by (20), for the torus it is based on numerical integration of (21) and for the box we resort to an iterative solution of the original quenched equations. The accuracy of this iterative technique is confirmed by reproducing the results for the other domains. In all these cases we have also performed Monte Carlo simulations and obtain excellent agreement with the theory. At large positive β the topology starts to become unimportant and each domain has average path length $\sim 2/\beta$ as expected for a two dimensional TSP.

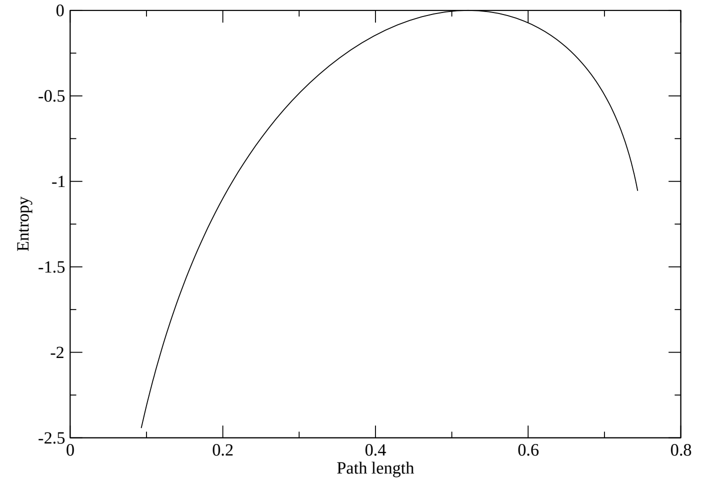

FIG. 2: Theoretical prediction for the entropy, the log of the number of paths, for two dimensional TSP on a box as a function of the path length.

IV. COUNTING PATHS

The partition function also contains information about the number of paths that contribute at given temperature or path length. We first compute the entropy as:

$$\frac{S(\beta)}{N} = \left(\beta E + \frac{\ln(Z_N)}{N}\right) = 2\left(1 - \beta\frac{\partial}{\partial\beta}\right)\left[\int d^d r \; \ln\left(s(\mathbf{r})\right)\right] \tag{22}$$

Then by inverting the expression for the path length obtained from equations (17) and (18), the entropy can be re-expressed as a function of path length. However, even for the closed domains discussed in the last section, this last step cannot be performed analytically. For the case of a two dimensional square domain, a numerical solution is shown in figure 2.

The significance of the entropy in this problem is that the number of paths of given length is:

$$\mathcal{N}(E) \sim N! \; e^{NS(E)} \tag{23}$$

Of course, as for our expressions for the path length, this N dependence arises from our way of solving the problem and does not correctly describe the region where path lengths scale as $N^{1-1/d}$. However in the case of the one dimensional wallpaper problem [12], where the number of paths is given by an Eulerian number, we can independently recover this form with $S(E)$ matching our results.

Most interest attaches to the shortest paths, and in this regime (though remaining with our N scaling), we can make some analytic progress. At least for closed domains leading to constant s, we can expand for large β to find $E \sim d/\beta$, and $S(\beta)/N \sim -d\ln(\beta)$, so we expect $\mathcal{N}(E) \sim N!\,(E/N)^d$.

V. ZERO TEMPERATURE LIMIT

In this section we investigate the low temperature behavior of Eqs. (17,18) and extract formulae for the average optimum path length. We write $s(\mathbf{r}) = \exp(-\beta w(\mathbf{r}))t(\mathbf{r})$ so Eq. (18) becomes

$$\exp\left(-\beta w(\mathbf{r})\right) t(\mathbf{r}) = \int d^d\mathbf{r}' \; \frac{\exp\left(-\beta|\mathbf{r}' - \mathbf{r}| + \beta w(\mathbf{r}')\right)}{t(\mathbf{r}')}. \tag{24}$$

We require that the decomposition in terms of $w(\mathbf{r})$ and $t(\mathbf{r})$ is such that $\ln(t(\mathbf{r}))/\beta \to 0$ as $|\beta| \to \infty$ permitting us to evaluate the integral on the right hand side of Eq (24) via the saddle point method as $|\beta| \to \infty$. The saddle point equation determines the function $w(\mathbf{r})$ through:

$$w(\mathbf{r}) = \min_{\mathbf{r}' \in \mathcal{D}} \left(|\mathbf{r}' - \mathbf{r}| - w(\mathbf{r}')\right) = \left(|\mathbf{r}' - \mathbf{r}| - w(\mathbf{r}')\right)_{\mathbf{r}'=\mathbf{r}^*} \tag{25}$$

where the min is replaced by max for the maximal (negative β) case.

The point $\mathbf{r}^*(\mathbf{r})$ is the point about which the action in the saddle point is minimal. It is tempting to suggest a tentative interpretation of $\mathbf{r}^*(\mathbf{r})$ in terms of an algorithm for finding the optimum. We take a local algorithm that selects cities one by one and determines the next city on the path solely on the basis of the location of the present city. This local class of algorithms is natural from a physical perspective but is severely restricted, and indeed does not include many well known algorithms for solving the minimal TSP, however for the leading large N behavior we are concerned with, it is sufficient. Then $\mathbf{r}^*(\mathbf{r})$ is interpreted as the location of the subsequent city to be chosen by the algorithm if the current city is at $\mathbf{r}$. In other words, if the i'th city is at $\mathbf{r}$, the $(i+1)$'th is at $\mathbf{r}^*(\mathbf{r})$.

Intriguingly, we will see that effectively guessing some (local) heuristics for $\mathbf{r}^*(\mathbf{r})$ will enable us to obtain some solutions to Eq. (25) and in [11], we have explored this approach for a variety of one-dimensional examples. For example, for the minimal TSP we take $\mathbf{r}^*(\mathbf{r}) = \mathbf{r}$ corresponding to picking the closest city. This algorithm has been known to be incorrect since the very earliest days of the TSP, however in the present context we only use it to demonstrate that the contribution to the path length proportional to N vanishes. To see this, use it to solve Eq. (25) finding $w(\mathbf{r}) = 0$. The leading term in the path length, according to Eq. (17) is:

$$E \approx 2 \int_{\mathcal{D}} d^d\mathbf{r}\; w(\mathbf{r}) \tag{26}$$

which therefore vanishes. A more detailed analysis taking proper account of the fluctuations in $t(\mathbf{r})$ gives the corrections to this term at small but finite temperature as a series in $1/\beta$. More interesting might be corrections for finite N at zero temperature as this would measure the first non-vanishing $O(N^{1-1/d})$ contribution to the path length, but we have not yet been able to address this with our formalism.

We therefore move to the MAX TSP. In this case we consider a greedy algorithm which selects the city most distant from the present one. In the case of the closed symmetric two dimensional domains, the torus and sphere, this next city is always at half the maximum period, $\mathbf{l}$, distance away: $\mathbf{r}^*(\mathbf{r}) = \mathbf{r} + \mathbf{l}/2$. We therefore have $w(\mathbf{r}) + w(\mathbf{r}+\mathbf{l}/2) = |\mathbf{l}|/2$, and because of the symmetry of the space $w(\mathbf{r})$ takes the constant value $|\mathbf{l}|/4$. Using Eq. (26) for the path length we regain the maximum average path length as half the maximum period which we earlier found from taking the low temperature limit of the full solutions in Eq. (20) and (21).

In the case of a unit hypercube $[-\frac{1}{2}, \frac{1}{2}]^d$ (it is convenient to shift the coordinates to be symmetric), consider the greedy heuristic; we start on the outermost layer of points in the hypercube and we join points on this surface to those that are diametrically opposed. The procedure is then repeated eating away the hypercube until we arrive at the center. This entails matching the point $\mathbf{r}$ with $-\mathbf{r}$. This heuristic was shown to give the optimal path length for $d = 2$ [13]. The solution $w(\mathbf{r}) = |\mathbf{r}|$ is in fact a solution to the maximal version of Eq. (25). This can be verified as we note that the function

$$h(\mathbf{r}) = |\mathbf{r} - \mathbf{r}'| - |\mathbf{r}'| \tag{27}$$

is bounded as

$$h(\mathbf{r}) \le |\mathbf{r}| \tag{28}$$

by the triangle inequality. The bound is achieved at $\mathbf{r}^*(\mathbf{r}) = -\mathbf{r}$, confirming that $w(\mathbf{r}) = |\mathbf{r}|$ is indeed a solution. We note that this solution exists in any domain $\mathcal{D}$ (centered at the origin) satisfying the property that if $\mathbf{r} \in \mathcal{D}$ then $-\mathbf{r} \in \mathcal{D}$. The ground state path length generated by this greedy heuristic is

$$E = \int_{[-\frac{1}{2},\frac{1}{2}]^d} d^d\mathbf{r}\; 2|\mathbf{r}| = 2^{d+1} \int_{[0,\frac{1}{2}]^d} d\mathbf{r}\; |\mathbf{r}|. \tag{29}$$

This general formula gives $E = 0.5,\ 0.765196,\ 0.960592$, in one, two and three dimensions respectively, in agreement with the results of [13].

VI. CONCLUSIONS

We have shown how a statistical mechanics approach can address the TSP. We have considered a stochastic Euclidean version of the problem and concentrated on computing the value of the average optimal path, but have also shown that the method can predict the number of paths of given length. The method sheds new light on the problem and it is that, rather than any specific results, that we wish to emphasize. In particular, it is fascinating that the solution to the zero temperature equations is illuminated by considering a local greedy algorithm.

The derivation we have presented relies heavily on singular distributions, and although results have been confirmed by simulations and various sanity checks, a more secure foundation (probably discrete) would be welcome. A significant reason to explore this direction would be a better understanding of finite N corrections to our formalism that we hope may have a bearing on the minimal TSP.

Besides the main concern with maximal and minimal versions of the Euclidean TSP in the traditional square domain, we would like to point out several interesting issues. Firstly, the closed symmetric domains such as sphere and torus in two dimensions have a significantly simpler solution (they are also expected to have simpler finite N corrections [4]). It is not clear that this is or should be reflected in heuristics. Secondly, our formalism is able to accommodate non-uniform distributions of cities and this may allow it to be applied to generalized TSP versions in which the cities cluster [2]. Thirdly, although the method is certainly geometric, being based on points in the domain $\mathcal{D}$ in $\mathbb{R}^d$, the norm need not be Euclidean. Indeed, in [11], we considered other norms, though we interpreted them as potentials between points on a closed polymer.

Acknowledgment: We would like to thank N. Read for drawing our attention to reference [13]. D.S.D. would like to acknowledge that part of this work was carried out at the KITP, UCSB and was supported in part by the National Science Foundation under Grant No. PHY99-0794. D.L would like to acknowledge a Discipline Hopping Award from the EPSRC.

[1] O.C. Martin, R. Monasson and R. Zecchina, Theoretical Computer Science **265**, 3 (2001).
[2] A. Barvinok, E. Kh. Gimadi and A.I. Serdyukov in G Gutin and, A. P. Punnen (Eds), The Traveling Salesman Problem and its Variations (Combinatorial Optimization Series), Kluwer, Boston (2002).
[3] J. Beardwood, J. H. Halton and J.M. Hammersley, Proc. Cam. Phil. Soc. **55** 299 (1959).
[4] D.S. Johnson, L.A. McGeoch, E.E. Rothberg, Proc. 7th ACM-SIAM Symposium on Discrete Algorithms, 341 (1996).
[5] J. Vannimenus and M. Mézard, J. Physique Lett. **45** L1145 (1984).
[6] M. Mézard and G. Parisi, Europhys. Lett. **2**, 913 (1986).
[7] W. Krauth and M. Mézard, Europhys. Lett. **8**, 213 (1989).
[8] A.G. Percus and O.C. Martin, Phys. Rev. Lett. **76**, 1188 (1996).
[9] A.G. Percus and O.C. Martin, J. Stat. Phys. **94**, 739 (1999).
[10] D.S. Dean, D. Lancaster and S.N. Majumdar, J. Stat. Mech. L01001 (2005).
[11] D.S. Dean, D. Lancaster and S.N. Majumdar, Phys. Rev. E **72** 026125 (2005).
[12] R.S. Garfinkel, Oper. Res. **25**, 741 (1977).
[13] M.E. Dyer, A.M. Frieze, and C.J.H. McDiarmid, Oper. Res. Lett. **33**, 267 (1984).