Configuration space analysis of travelling salesman problems

S. Kirkpatrick, G. Toulouse

J. Physique **46** (1985) 1277-1292 AOÛT **1985**, PAGE 1277

Classification
Physics Abstracts
05.20 — 75.50 — 85.40

Configuration space analysis of travelling salesman problems

S. Kirkpatrick

IBM T. J. Watson Research Center, Yorktown Heights, N.Y. 10598, U.S.A.

and G. Toulouse (*)

Laboratoire de Physique de l'E.N.S., 24, rue Lhomond, 75231 Paris Cedex 05, France

(*Reçu le 14 février 1985, accepté le 18 avril 1985*)

Résumé. — Le problème du voyageur de commerce (TSP) et le modèle d'Ising d'un verre de spin sont, respectivement, des archétypes pour les problèmes d'optimisation combinatoire en informatique et pour les systèmes désordonnés frustrés en physique de la matière condensée. Il a été suggéré récemment que ces deux domaines ont beaucoup de phénomènes en commun. Pour voir si, de fait, les problèmes d'optimisation combinatoire peuvent être des verres de spin, nous définissons un TSP à distance aléatoire aussi semblable que possible au modèle idéalisé, à portée infinie, des verres de spin. A la lumière des résultats récents pour les verres de spin, nous analysons les observables thermodynamiques et les corrélations internes entre configurations localement stables. L'hypothèse d'un gel dû à la frustration et d'une structure hiérarchique ultramétrique dans l'espace des configurations est solidement argumentée, pour ce problème de voyageur de commerce.

Abstract. — The travelling salesman problem (TSP) and the Ising model of a spin glass are archetypes, respectively, of the combinatorial optimization problems of computer science and of the frustrated disordered systems studied in condensed matter physics. It has recently been proposed that these two fields have many phenomena in common. To see if, in fact, combinatorial optimization problems may be spin glasses, we define a random distance TSP as similar as possible to the idealized infinite-ranged model of spin glasses. Thermodynamic observables and internal correlations among locally stable configurations are analysed in the light of recent results for spin glasses. Evidence for freezing due to frustration and for a hierarchical, ultrametric structure of configuration space in this TSP is presented.

The travelling salesman problem (TSP) is one classic example of a complex optimization problem. Easy to formulate, it is yet hard to solve : it belongs to the class of *NP*-complete problems [1]. Recently, the use of simulated annealing, a stochastic algorithm based on the Monte Carlo method (as developed in the context of the statistical physics of disordered systems, e.g. spin glasses), was advocated as an appropriate algorithm for approximate solution of complex optimization problems [2]. The application of the simulated annealing method to TSP has already been discussed to some extent [2, 8].

There are good reasons why spin glasses provide a suggestive model for complex optimization problems [2]. They possess, in a clear way, the physical ingredients of frustration and disorder, which lead to a large number of locally minimal solutions and to freezing phenomena. The simulated annealing method addresses the problem of getting stuck in a local minimum by allowing for uphill moves.

Meanwhile, new advances in the theory of spin glasses have directed attention toward the distribution of local minima in the landscape of configuration space. Some sharp results have been obtained for long range Ising spin glasses : ultrametricity, non-reproducibility, loss of self averaging [9, 10, 23]. Several new concepts and tools have thus been created.

(*) Present address : ESPCI, 10 rue Vauquelin, 75231 Paris, France.

The notion of configuration space landscape is obviously so general as to be of interest in many problems. This paper is a study, largely numerical, of the characteristics of configuration space landscapes for a random distance version of the TSP.

Section 1 describes the problem and our analysis.

Section 2 presents the numerical data.

Section 3 surveys the recent relevant results obtained in the theory of spin glasses, and draws some comparisons with our TSP data.

Some general views can be extracted from these case studies. The analysis of configuration space landscapes aims at finding a more detailed and more physical categorization among complex optimization problems. Numerical and analytical strategies for solving such problems will thus be based on sounder physical grounds. (Present problem classifications, based on worst case analysis of algorithmic complexity, are often very broad, so for instance the problem of finding the ground state of a ferromagnet might be considered *NP*-complete because it is lumped together with the spin glass problem.)

Finally, in section 4, we discuss the impact of these ideas on the monitoring of the change of configuration space landscapes in memory models, under learning and unlearning processes.

1. Introductory analysis.

The travelling salesman problem (TSP) is simply stated. A list of N cities and a means of obtaining the distance between any pair of cities is given. The objective is to find a tour, or permutation, P, of the cities such that the total length L, of travel through all cities and returning to the first one in the order P:

$$L = \sum_{i=1}^{N} d_{P(i)P(i+1)}, \quad \text{where} \quad P(N+1) \equiv P(1), \quad (1)$$

is minimized. Since the starting point and direction of the tour do not matter, there are $(N-1)!/2$ distinct tours. Heuristic strategies which search for near-optimal tours will be able to explore only a tiny fraction of this enormous configuration space if N is large.

S. Lin [11] has introduced a natural rearrangement of a given tour which permits efficient search. Two steps on the tour are discarded, say $d_{i,i+1}$ and $d_{j,j+1}$. They are replaced by $d_{i,j}$ and $d_{i+1,j+1}$ so that the new path is again a tour. There are $N(N-1)/2$ such moves from any tour.

« Iterative improvement », or exhaustive search using this class of moves until no further improvements can be found, is a powerful way of improving upon an initial guess of a reasonable tour. Several groups have recently [2-7] used the Metropolis algorithm to enhance the effectiveness of search with these moves. In general they find solutions with annealing and Lin's 2-bond moves which are as good as can be found with exhaustive search using replacement of 3 or more bonds.

The 2-bond moves impose a natural topology on the configuration space, giving each tour $O(N^2)$ neighbours. A tour which is shorter than all of its neighbours is termed « 2-optimal » or 2-opt. This is not the only possible notion of neighbourhood. For instance, Černý [4], Lundy [6], and Kirkpatrick (unpublished) have attempted searches using the N interchanges of successive cities on the tour or the N^2 interchanges of any pair of cities, but these moves usually turn short configurations into long ones, and do not lead to effective searches.

A fast, but more limited method of sampling the configuration space should also be mentioned. Called a « greedy algorithm » because it has finished in only N steps it is a useful means of obtaining trial configurations. One picks one city at random, takes the closest city as the second step, the closest remaining city as the third, and so on until the tour is complete. This differs from an iterative search in that at most N configurations can be constructed (in practice, fewer than N distinct tours are usually found).

The nature of the distances d_{ij} has a big influence on the difficulty of the TSP. In practical problems the d_{ij} may be calculated in an appropriate metric from the positions of the N cities, or they may be tabulated quantities, representing some more complicated cost. For N points randomly distributed in a d-dimensional Euclidean unit cube, Beardwood *et al.* [12] have shown that the minimal tour length is proportional to N times the expected nearest neighbour distance, $N^{-1/d}$. In this case, all the heuristics discussed approach the exact limit and give expected outcomes proportional to $N^{(1-1/d)}$. The greedy algorithm gives configurations 20 percent longer than optimal in some 2-dimensional problems [2] (KGV), while the other heuristics are within a few percent of each other.

In this paper we shall consider a simple but artificial TSP in which the distances are symmetric ($d_{ij} = d_{ji}$) random variables independently drawn from the uniform distribution over the interval (0, 1).

This model is appealing for its simplicity and freedom from geometry. We hope that it may eventually prove, as has the S. K. model of spin glasses [13], to be analytically tractable and provide a « mean field » limit of the statistical mechanics of a travelling salesman. However, Vannimenus and Mézard [14] have pointed out that the uniform distribution is not the infinite dimensional limit of the distribution of distances between points placed at random in a finite d-dimensional volume. Still, there is inherent simplicity in the uniform distribution which explains our choice.

The expected nearest neighbour distance is the smallest of $(N-1)$ random numbers in (0, 1), hence $O(N^{-1})$, in this model. Thus an N-step tour must be at least 1 unit long, and may, as an extension of the Beardwood result, have expected length tending to

be constant O(1) as $N \to \infty$, but we know of no proof of this. The greedy algorithm for this random distance problem gives expected length

$$L_{\text{greedy}} \sim (N-1)^{-1} + (N-2)^{-1} + (N-3)^{-1} + \cdots \sim \ln N . \quad (2)$$

and provides an upper bound.

In figure 1 we display tour lengths obtained for this random distance TSP by exhaustive 2-opt and 3-opt searches, and by simulated annealing based on 2-bond and limited 3-bond moves. This problem is difficult enough to discriminate asymptotically between different heuristics. As with 2-dimensional TSP's, it was found that simulated annealing with 2-bond moves gave answers as good as exhaustive search with 3-bond moves for comparable computing times. In figure 1, the annealing schedule was adjusted so that the computing time for the annealed runs was kept $O(N^2)$. Thus, exhaustive 3-opt, with computing time $\propto N^3$, became better than 2-bond annealing for large N. For $N \geqslant 100$, it was extremely costly to obtain good configurations by any heurisic : for $N \leqslant 50$, the solutions appear to be close to the presumed optima.

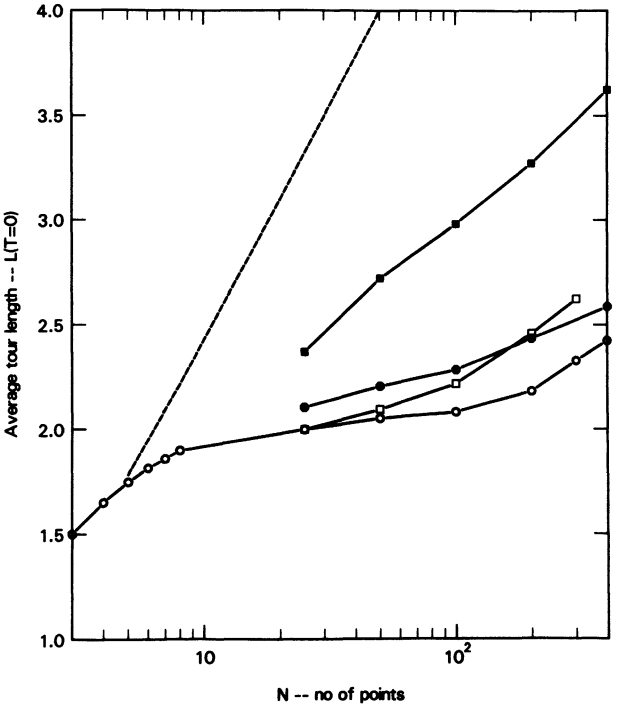


Fig. 1. — Optimal tour length for random distance TSP's of up to $N = 400$ sites, as obtained by various algorithms. The dashed line is the upper bound provided by the greedy algorithm. The solid points are from iterative improvement using exhaustive search for 2-bond rearrangements (squares) and 3-bond rearrangements (dots). The open points for $N \geqslant 24$ are simulated annealing results using 2-bond moves (squares) and three bond moves (circles). The open data for $N \leqslant 12$ are exact results.

Some simple calculations will suffice to sketch in the important energies (i.e. lengths) and temperatures of the random distance TSP, viewed as a statistical mechanics problem. At high temperatures we neglect the fact that the configuration must be a tour and simply treat each bond as contributing with a Boltzmann weight. This corresponds to the « annealed » rather than « quenched » approximation used in disordered systems. It is known to give a useful description of spin glasses above freezing, and was applied in an independent study by Bonomi and Lutton [5] to describe a TSP in 2D at non-zero temperature.

The statistical mechanics of this « annealed » approximation follows simply from treating each of the N bonds as independent. The partition function, Z, is approximated by

$$Z(\beta) \sim N!\, g(\beta)^N , \quad (3)$$

where

$$\beta = 1/T$$

and

$$g(\beta) = \int_0^1 dl\, e^{-\beta l} = (1 - e^{-\beta})/\beta . \quad (4)$$

The path length per bond, $l(\beta) = L/N$, is obtained by

$$l(\beta) = -\frac{\partial}{\partial \beta} \ln g(\beta) \quad (5)$$

$$= \beta^{-1} - (e^{\beta} - 1)^{-1} . \quad (6)$$

This has limits $l \sim 1/2$ at high temperature and $l \sim T$ at low temperature, with a crossover around $\beta \sim 1$. But, since $l \sim N^{-1}$ is the smallest possible for actual tours, the annealed theory will have broken down for $\beta \sim N$.

The length per bond, $l(\beta)$, plays the role of internal energy in the TSP, so the specific heat, $C(\beta)$, is obtained by

$$C(\beta) = \frac{\partial}{\partial T} l(\beta) = 1 - \beta^2 e^{\beta}/(e^{\beta} - 1)^2 , \quad (7)$$

which has limits

$$C \sim 1 - \beta^2 e^{-\beta} \quad \text{as} \quad \beta \to \infty , \quad (8)$$

and

$$C \sim \frac{\beta^2}{12} \quad \text{as} \quad \beta \to 0 . \quad (9)$$

The entropy, S, can be obtained by integrating C/T, or by defining a free energy, F,

$$F(\beta) = -T \ln Z(\beta) \simeq NT \ln \frac{N}{e} g(\beta) . \quad (10)$$

Now the entropy per step of the tour is given by

$$S/N = N^{-1}\, \partial F(\beta)/\partial T ,$$

with the result

$$S/N = 1 - \beta/(e^{\beta} - 1) + \ln\left[\frac{N}{e}(1 - e^{-\beta})/\beta\right]. \tag{11}$$

At high temperatures,

$$S/N \sim \ln\frac{N}{e} - (24\,T^2)^{-1}, \tag{12}$$

and at low temperatures,

$$S/N \sim \ln(NT). \tag{13}$$

This negative entropy catastrophe is familar from classical statistical mechanics. Since negative entropy is unacceptable for a problem with discrete states, it gives a strong limit for the validity of the annealed model,

$$S/N > 0 \quad \text{when} \quad \beta < N. \tag{14}$$

We plot the predictions of the annealed model in figures 2a-b for $N = 48$, and compare with Monte Carlo data. The data are in good agreement with the annealed theory for temperatures above $\sim 2/N$. Below this temperature the average bond length begins to saturate due to the frustration of being required to complete a tour, and the simulations exhibit all the usual phenomena of freezing. Similar observations have been made by the groups who have studied the TSP in 2 dimensions at finite temperatures [5].

Because the total length L of a tour is of order N at high temperature, while it is of order 1 at zero temperature, a phase transition between these two different scaling behaviours must occur. The data of figure 2 and other data, not represented here, for a number of values of N do not provide evidence for a phase transition occurring at finite temperatures. Therefore, in the absence of an analytical solution of the problem, the most conservative guess is that the phase transition occurs at zero temperature. The deviations from the annealed approximation, on figure 2, can be interpreted in terms of crossover behaviour between the $T = 0$ regime and the high temperature regime. But certainly it must be kept in mind that the spin glass phase transition, in the presence of an applied field, for the infinite range model, is too subtle to be observed numerically : its existence is known only from analytical calculations. Thus the question of the phase transition for TSP remains open.

On the other hand, abundant evidence for freezing is found and a principal goal in this study is to see if this freezing transition in the simplest TSP presents any of the characters which have been found for the freezing transition in spin glass models : self-similar structure of the configuration space landscape, fluctuations from valley to valley in a given

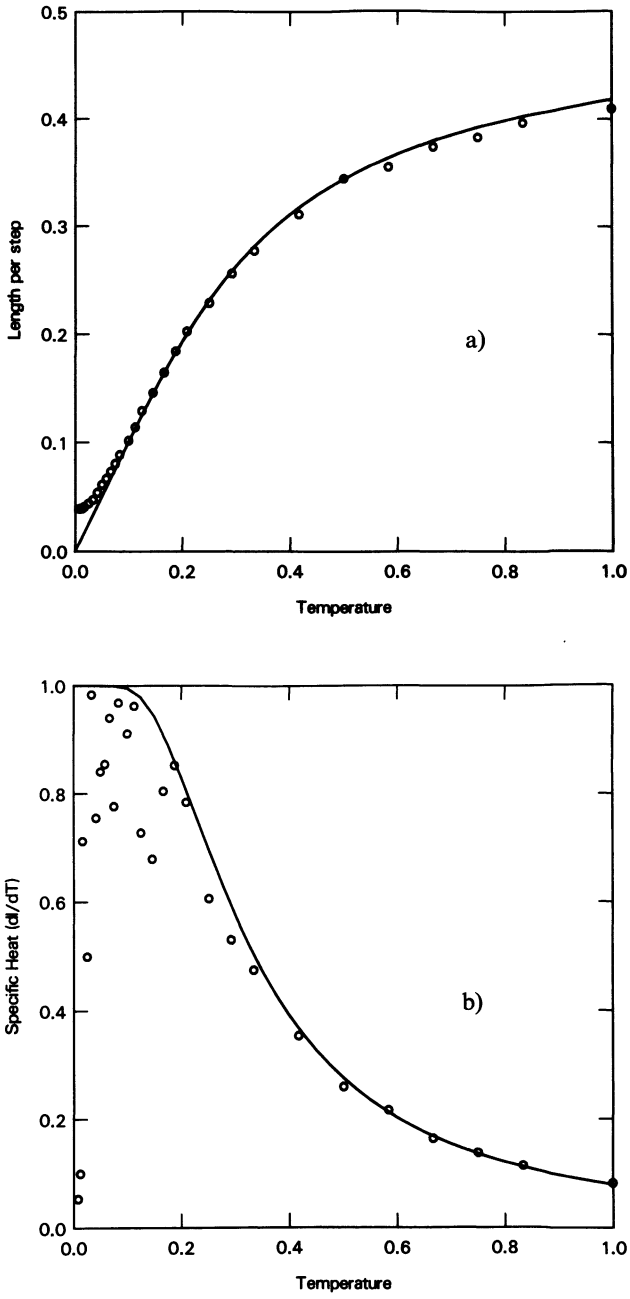


Fig. 2. — Static thermodynamic quantities for a random distance TSP of 48 cities. Shown are expected length per step (a), and specific heat (b), as calculated in the high temperature « annealed » approximation and as found experimentally by Monte Carlo simulation.

sample, fluctuations from instance to instance. Since we do not have yet a soluble mean field theory for TSP, we shall draw inspiration from the spin glass theory which is built on an analysis of overlaps between configurations. Here, we define the overlap, q_{ij}, between configurations i and j, to be the fraction of bonds which are common to both tours, without regard for the direction in which the bonds are traversed. This definition is consistent with our choice

of 2-bond moves to generate a topology in the configuration space since it implies that two configurations are close when only a few such moves are applied to the first to generate the second.

Some annealed results are of interest in thinking about the overlap of tours. Two randomly selected tours of N cities will have, on average, 2 bonds in common, and the distribution, $P(q)$, of overlaps is Poisson in the limit of large N [15].

If we include the effect of temperature as in the previous discussion, we find the expected overlap is given by

$$\langle Q_2(\beta) \rangle = (2/N)\, g(2\beta)/(g(\beta))^2 \tag{15}$$

for pairs of tours, weighted by their equilibrium probability of occurrence.

For p tours, the comparable expression is

$$\langle Q_p(\beta) \rangle = (2/N)^{p-1}\, g(p\beta)/[g(\beta)]^p \tag{16}$$

$$\approx [2\beta/N(1 - e^{-\beta})]^{p-1} \frac{1 - e^{-\beta p}}{p(1 - e^{-\beta})}. \tag{17}$$

From (17) we see that little overlap is expected when $\beta \approx 1.0$, and significant overlap once $\beta \approx N$, where we have shown the annealed theory is inadequate.

At low temperatures, when there is significant overlap between tours, we would like to know more about the bonds which participate in large numbers of tours. Questions like how frequently a given bond appears in the sample, or how many bonds take part in a given class of locally optimal states are of interest.

For a set $\{ \alpha \}$ of M tours (for example, all of the 2-opt configurations), define

$$n_i^\alpha = 1 \text{ if bond } i \text{ is in tour } \alpha\,,$$
$$= 0 \text{ if it is not}\,. \tag{18}$$

Then

$$N_i = \sum_\alpha n_i^\alpha \tag{19}$$

is the number of occurrences of the i-th bond. Its relative frequency $f_i \equiv N_i/M$. The overlap between p tours, $Q^{\alpha_1,\ldots,\alpha_p}$ can be expressed as

$$Q^{\alpha_1,\ldots,\alpha_p} = \frac{1}{N} \sum_i \prod_j n_i^{\alpha_j} \tag{20}$$

so

$$\langle Q_p \rangle = \frac{1}{N} \sum_i (N_i/M)^p\,. \tag{21}$$

Thus the frequency of bond occurrence can be used to generate higher order overlap statistics.

Now if we introduce a frequency distribution, $p(f)$, normalized such that

$$\int_0^1 p(f)\, \mathrm{d}f = N \frac{(N-1)}{2}, \tag{22}$$

the identity

$$\sum_i \frac{N_i}{M} = N \tag{23}$$

gives a useful sum rule on $p(f)$:

$$\int_0^1 f p(f)\, \mathrm{d}f = N\,. \tag{24}$$

We show numerical evidence in the following section that for the set of 2-opt configurations, $p(f)$ consists of a nearly constant piece plus a delta function at $f = 0$, representing the bonds which never participate. Introducing an ansatz, $p_0(f)$, for the density with this two-part form :

$$p_0(f) = A f^{-\gamma} + \left(\frac{N(N-1)}{2} - \frac{A}{1-\gamma} \right) \delta(f)\,, \tag{25}$$

we find from (25) that $A = (2 - \gamma)\, N$ and as a result the number of bonds which participate in the sample of M configurations will be $\left(\frac{2-\gamma}{1-\gamma} \right) N$. This form for $p_0(f)$ implies that only a finite number of bonds per city participate in the locally stable solutions of a TSP, a hypothesis which suggests intriguing heuristics for restricting the search for an optimal solution of the TSP.

There are additional quantities of interest in studying TSP, but since we have no theoretical estimates these will be introduced below, during the discussion of the numerical data.

2. Numerical results.

The numerical work in this paper is based upon populations of distinct locally minimal tours which were extracted from several samples of the random distance TSP in each of three sizes : $N = 12$, 24, and 48. Exact enumeration was employed to study tours of 12 cities. Monte Carlo methods were required for $N = 24$ and above. TSP's with $N = 96$ were also studied but are not reported because the tours obtained were probably not close to optimal, as figure 1 suggests.

Eleven instances, i.e. matrices of random distances d_{ij}, were analysed for $N = 12$. For each instance, all 2-opt configurations were extracted and saved in a disk file. There are approximately 4×10^7 distinct tours of length 12. Generating all tours and calculating their lengths required about 4 min on an IBM 3081. We used an algorithm due to Trotter [16] to generate permutations by N ! successive interchanges of adjacent elements. Trotter's algorithm appears to be optimal for our purposes. The 11 minimal tours obtained ranged in length from 1.16 to 2.47, with a mean of 1.83 and variance of 0.44. From 11 to 58 2-opt tours were extracted; on average there were 29 per configuration. The lengths of the 2-opt tours ranged up to 3.3. The average spread in length of the 2-opt

states for a given instance was 0.9, the variance, 0.4. Since the mean length of a tour for $N = 12$ in this model is 6 and the maximum length O(12), the 2-opt tours are a more restricted class of states than the locally minimal spin glass configurations, which are found at energies up to the average energy of a spin configuration [17, 18]. A 2-opt solution is stable against $O(N^2)$ bond rearrangements, while a spin configuration with all spins frozen is stable against N individual spin reversals. Thus the narrower relative range of lengths for the 2-opt solutions is not surprising.

The 2-opt states were also tested for 3-optimality. On average, two 3-opt tours per instance were found, with some instances having a unique 3-opt configuration and others as many as 4.

To obtain 2-opt tours with $N = 24$ and 48, a simulated annealing procedure (KGV) was employed to generate 100 2-opt configurations for each instance, heating the system to a temperature of 2 or more briefly between coolings to provide a random restart. For $N = 24$, if the samples were cooled slowly, only a few distinct 2-opt states resulted. Thus a relatively rapid annealing schedule was developed, which gave an average of 55 distinct 2-opt tours in the set of 100 retained for each instance. One or two of the shortest 2-opt tours were generated 10 or more times each, while most of the longer 2-opt tours created were unique. We found from 3 to 23 of these tours were also 3-opt, for an average of 11 per instance. The range of lengths sampled is not relevant for $N = 24$, since the selection is not exhaustive or random. The same annealing schedule was followed in generating tours for four instances with $N = 48$, taking four times as many rearrangement steps at each temperature. Each cooling run terminated only when a 3-opt configuration was formed, which proved to occur frequently. The 100 3-opt tours generated for each instance were all distinct.

Next we show the distribution of pairwise overlaps found in the sample populations of 2-opt (or 3-opt) tours. Figures 3a-c show similar-looking curves for $P(q)$ with $N = 12$, 24 and 48, respectively. Curves for the different instances at each N are superimposed. In each case the most probable values of q occur at about 2/3. Note the differences between figures 3a-c and the overlap distributions predicted and found for spin glasses. There $P(q)$ is nonzero down to $q = 0$; overlaps of negative sign are meaningful between spin configurations. The distributions in figures 3 do not include self overlap, i.e. terms q_{ij} with $i = j$, and thus are normalized to slightly less than unity. However we prefer to avoid the delta functions this term requires and will also leave out self overlaps in the joint distributions explored below.

With increased temperature, the tours spread further apart in configuration space, reducing the average overlap, but the spread in $P(q)$ does not change much. Figure 4a shows distributions, $P(q)$, obtained for one

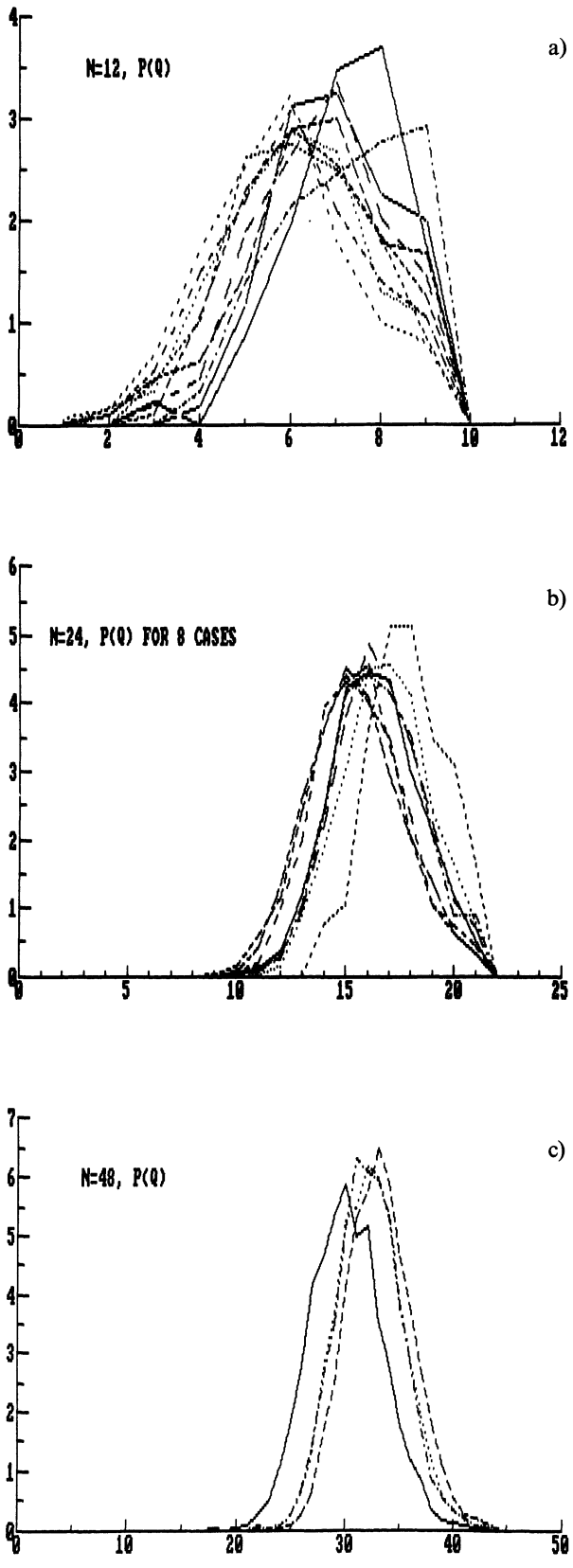
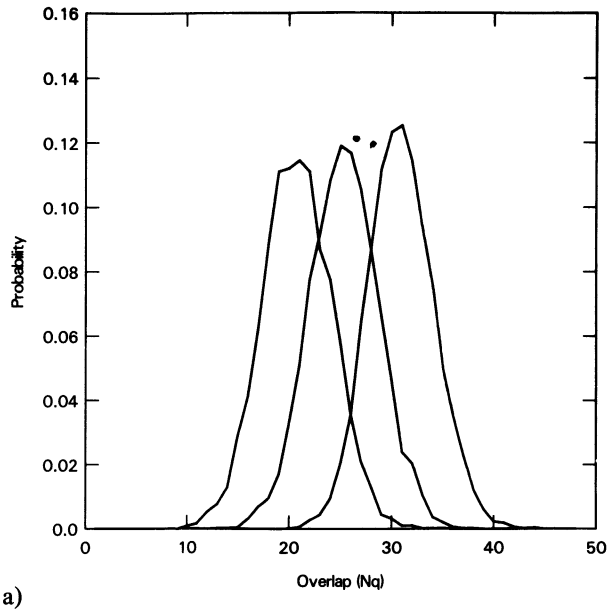
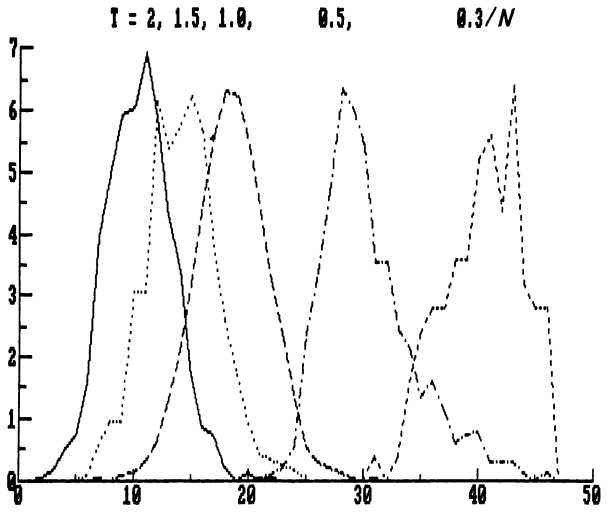


Fig. 3. — Overlap distributions, $P(q)$, for eleven instances of $N = 12$ (a), eight cases of $N = 24$ (b), and four cases with $N = 48$ (c). In figures 3, 4, 9-13, we plot data against the integervalued Nq for programming convenience. The vertical scales are arbitrary ; $P(q)$ is a normalized probability distribution.



a)



b)

Fig. 4. — $P(q)$ for samples annealed to temperatures 1/12, 1/24, and 1/48. In (a) random restarts were employed after each sample configuration was collected. In (b) the configurations were obtained by Monte Carlo runs at the indicated fixed temperatures, without warming up between samples. See the caption of figure 3 for notations.

instance with $N = 48$ in populations of 100 tours obtained by repeated cooling from the hot, scrambled state to various lower temperatures.

The effect of the energy landscape which confines the Monte Carlo search to a particular valley or to the vicinity of a set of minima is evident in figure 4b. Here we plot $P(q)$ obtained at a similar set of temperatures, but this time without randomizing the tour between samples. The final set of tours obtained overlap more strongly, with a most probable value of $q \sim (7/8)$, than do the tours found by repeated random restarts in figure 3c.

We have also studied the bond frequency distributions characterizing the samples of locally optimal configurations. The well-behaved experimental quantity which is closely related to $p(f)$ is the cumulative distribution, $D(f)$:

$$D(f) = \int_f^1 p(f')\, \mathrm{d}f' . \tag{26}$$

Figures 5a-c plot f as a function of $D(f)$ for the instances with 12, 24 and 48 cities, respectively. The abscissa and ordinate are interchanged for convenience in generating these figures, which were created by sorting the list of all bonds in each instance in decreasing order by frequency, then plotting bond frequency against that bond's position in the sorted list. If $p(f)$ were constant ($\gamma = 0$ in the ansatz p_0), then $D(f)$ would be a straight line, and only $2N$ bonds would participate in all 2-opt configurations. Furthermore, a constant bond frequency distribution $p(f)$ implies $\langle Q_p \rangle = \dfrac{2}{p+1}$, where $\langle Q_p \rangle$ has been defined in (21), which is in good agreement with the estimate $\langle Q_2 \rangle \simeq \frac{2}{3}$, obtained from the date of figure 3. In fact, figures 5 suggest that $p(f)$ is roughly constant, with some extra contributions close to $f = 0$ and $f = 1$. About $3N$ bonds in each instance contribute to our samples of optimal configurations. The distributions shown in figures 5 sharpen up as N increases, with the case $N = 48$ showing little instance-to-instance scatter.

Figure 6, based on the samples used in figure 5c, shows the effect of higher temperature in reducing configuration overlap. In figure 6 no bond occurs in all samples, and a larger number of bonds participate in the configurations analysed. This is consistent with the lower average value of q observed.

It seems of interest to ask whether other sets of configurations, more easily calculated, would contain the same set of useful bonds as are present in the annealed 3-opt configurations, which are relatively costly to compute. We consider the set of N « greedy » solutions to a given TSP obtained by starting at each city. (In fact, for $N = 48$, one or two of the greedy solutions is duplicated, but we did not bother to remove duplicates). The best greedy solutions found for $N = 48$ are about 50 percent longer than the best annealed solutions (see Fig. 1). The distribution of bond frequencies for the greedy set of configurations, shown in figure 7 for one $N = 48$ instance, is similar in shape to the curves of figure 5c, except for a tail extending out to $4N$ bonds. The rather long bonds taken at the end of a greedy tour, when a few scattered cities remain to travel between, could account for this tail.

There are other differences. In figure 8 we plot bond frequencies against bond length for greedy and annealed configurations of the same TSP. The annealed solutions consist of bonds with a narrow

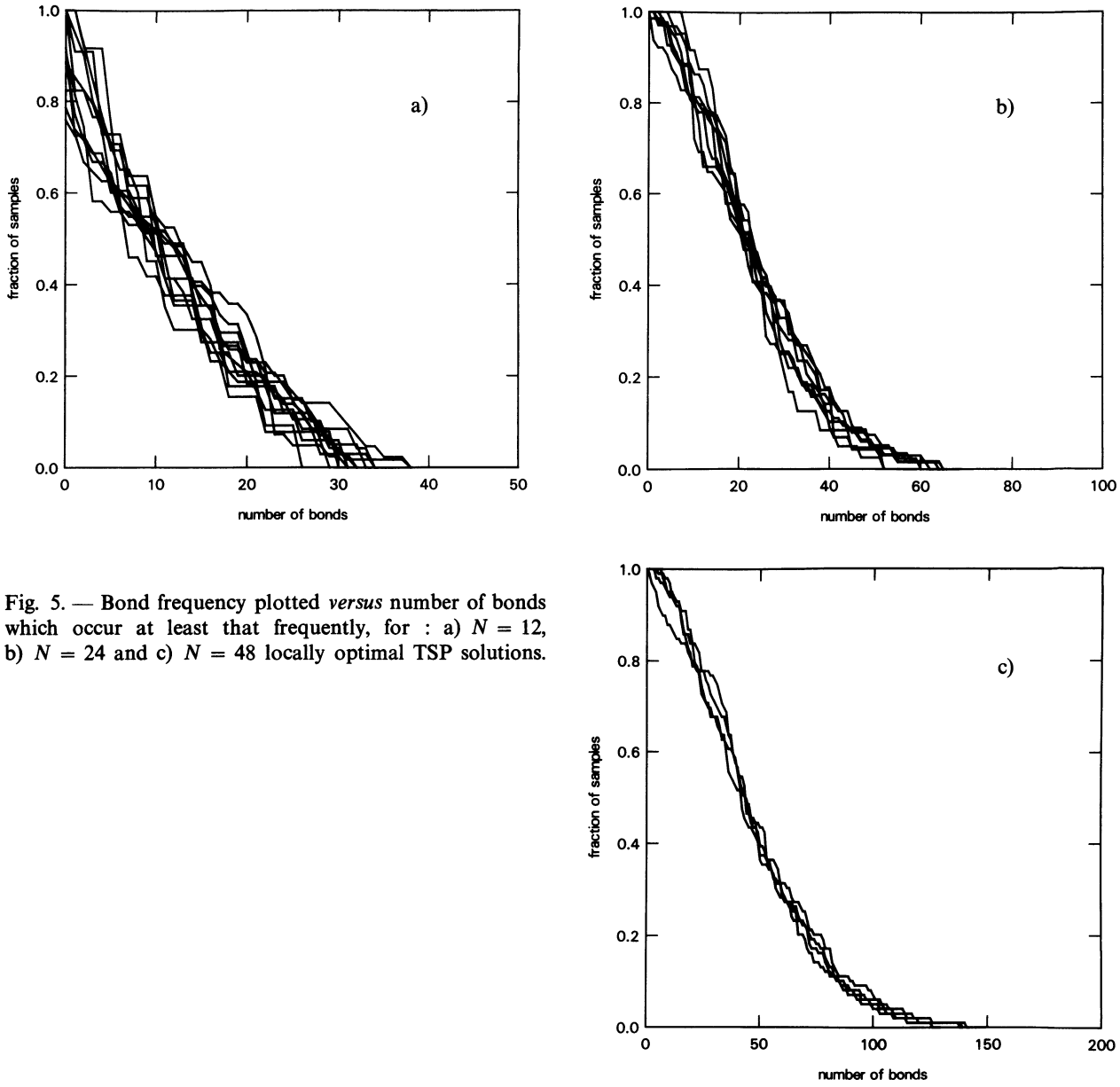


Fig. 5. — Bond frequency plotted *versus* number of bonds which occur at least that frequently, for : a) $N = 12$, b) $N = 24$ and c) $N = 48$ locally optimal TSP solutions.

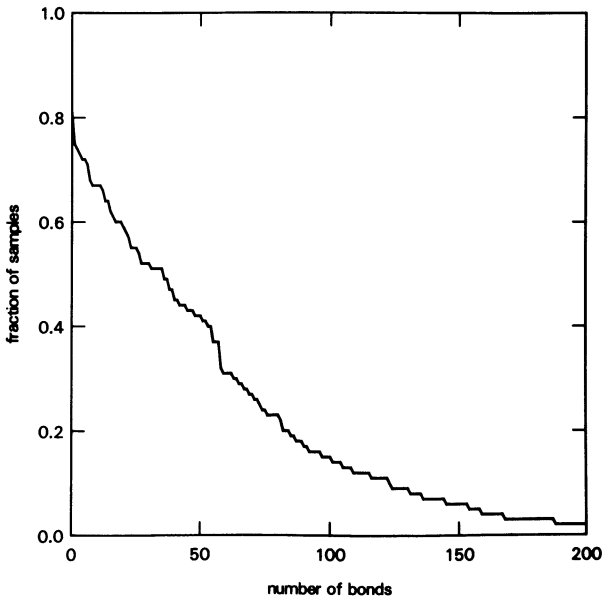


Fig. 6. — Cumulative distribution of bond frequencies, as in figure 5, for an $N = 48$ sample annealed to $T = 1/6$, then quenched.

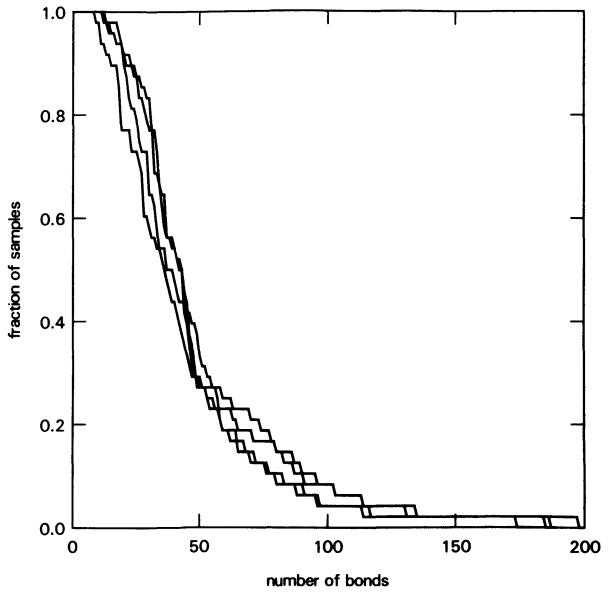


Fig. 7. — Cumulative distribution of bond frequencies for all greedy solutions of the same instances of TSP with $N = 48$ as in figure 5c.

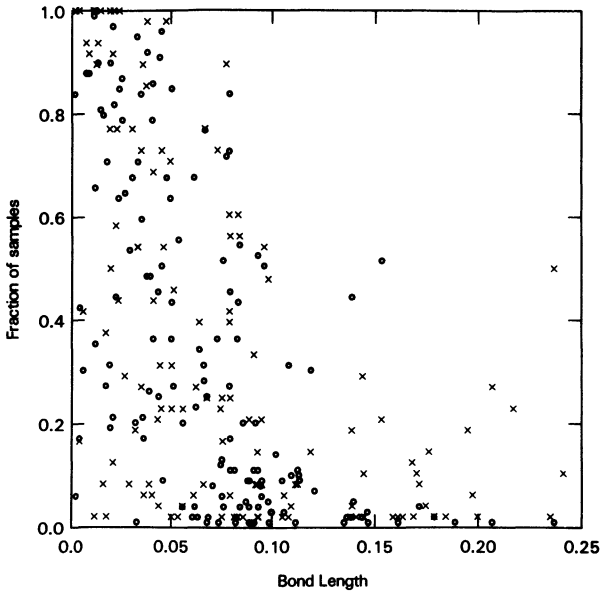


Fig. 8. — Relative bond frequencies plotted against bond length for 100 3-opt solutions (circles) and all greedy solutions (crosses) of an $N = 48$ TSP.

range of lengths, while some of the frequently encountered greedy bonds are quite long. Presumably, the rigid greedy strategy for creating a tour forces the same mistake repeatedly. Nonetheless, of the 60 shortest bonds in the annealed sample only three are not observed in the greedy sample. However, there are 37 bonds out of the 141 distinct bonds in the annealed sample of 100 configurations which were never generated by the greedy algorithm. Even if we concentrate on the 5 shortest tours obtained by annealing we find that these employ 76 distinct bonds, and six of them are not obtained in the greedy sample. Three of these six bonds occur in four or more of the five shortest tours. A sample of 50 tours was generated by cooling to a temperature at which freezing begins, $1.4/N$, using 50 independent random starts. This sample contained 192 distinct bonds, including all the bonds discovered in the 100 annealed near-optimal configurations.

Next we shall analyse the data for evidence of the ultrametric structure known to exist in spin glasses.

In a Euclidean space, the triangular inequality says that any side of a triangle is smaller or equal to the sum of the other two :

$$d(l_{ij}) \leqslant d(l_{ik}) + d(l_{kj}) . \qquad (27)$$

An ultrametric space is defined by a stronger inequality, the ultrametric inequality :

$$d(l_{ij}) \leqslant \text{Max} \{ d(l_{ik}), d(l_{kj}) \} . \qquad (28)$$

This implies that any triangle is either isosceles with small basis, or equilateral. Obviously, the preceding inequalities can be expressed in terms of overlaps, rather than distances. The ultrametric inequality was first put forward by M. Krasner, around 1930, in the context of arithmetics [19]. Notions of ultrametricity have also been used in the classification of data in various sciences, and especially in biological taxonomy [20]. There are various ways to define overlaps or distances between two biological objects, according to the characteristics that are under study. If the distance matrix of a collection of biological objects is ultrametric, then there is a unique, well defined, evolution tree such that the distance between two objects is determined by the closest common ancestor (the older the bifurcation the larger the evolved divergence, and thus the distance). In the general case, where the distance matrix is not purely ultrametric, biologists still wish to reconstitute an evolution tree. The two simplest procedures to do so are called « single linkage clustering » and « complete linkage clustering » [21]. In both procedures, the distance matrix is modified systematically to become ultrametric, with as few modifications as possible. In the first procedure, all distances are either decreased or unmodified, and the solution is unique. In the second procedure, distances are either increased or unmodified, and the solution need not be unique. So, these procedures somehow bracket the problem between ultrametric upper and lower bounds. If the evolution trees so obtained are analogous, this is an indication that the problem possessed a strong underlying ultrametric structure. If not, one has to resort to less stringent and more quantitative ultrametricity tests.

The triangle inequality (27) takes a slightly different form when the variables considered are overlaps, not distances. If we know that tours 1 and 2 differ in

only $N(1 - q_{12})$ of their bonds, and calculate overlaps q_{13} and q_{23} to a third state, then the maximum difference between q_{13} and q_{23} is

$$| q_{13} - q_{23} | \leqslant 1 - q_{12} . \quad (29)$$

The most direct check for ultrametric statistics is the distribution of the number of triangles (q_{12}, q_{23}, q_{13}) as a function of the lengths of the longer legs q_{23} and q_{13}. The triangle inequality may impose severe constraints on this plot. For example, see figure 9a, for $N = 24$, $Nq_{12} = 21$, where the resulting distribution is cut off at $N | q_{13} - q_{23} | = 4$. Part of the peaking in figure 9a is due to the narrow range of $P(q)$. We subtract the uncorrelated background $P(q_{13})\, P(q_{23})$ from the data in figure 9a to remove this effect, and show the result as figure 9b. While the removal of background leaves the ridge narrower and still rather high the negative contribution on either side of the ridge in figure 9b is deepest at the triangle constraint value $N | q_{13} - q_{23} | = 4$. For this reason, plots of the statistics of triangles are not informative about $N = 12$.

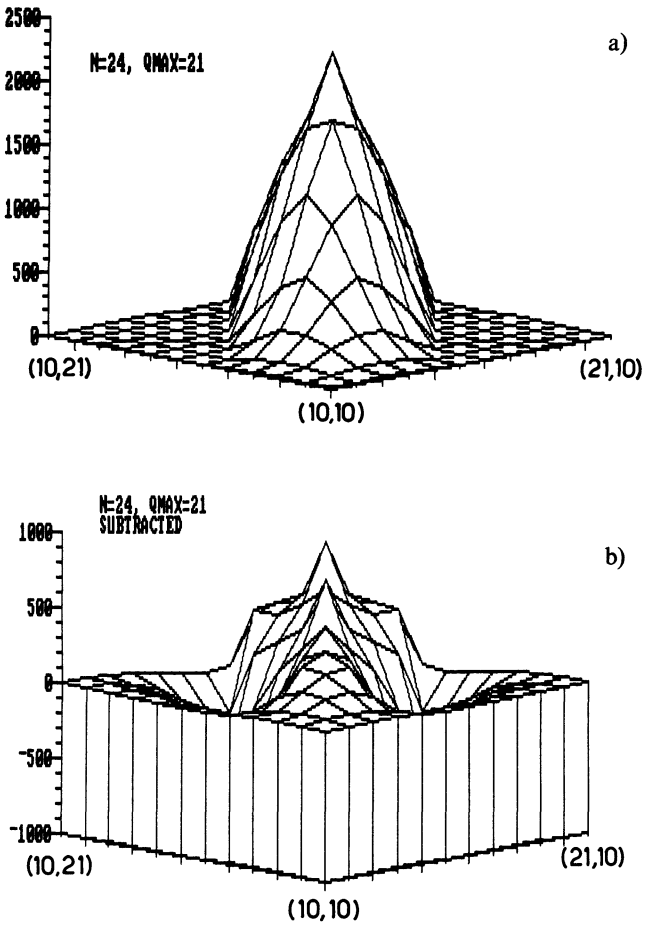


Fig. 9. — Distribution of triples of overlaps $q_{ij}\, q_{jk}\, q_{ki}$ for $N = 24$ for which the largest overlap is 21 bonds. In (a) is shown the raw data. In (b) the uncorrelated background is subtracted out. Vertical scales are arbitrary. The two horizontal variables are overlaps.

The triangle data is not as compromised for the case $N = 48$. Data for $Nq_{12} = 33$, the most probable value of q, are plotted with the uncorrelated background subtracted out in figure 10a. The restriction $N | q_{13} - q_{23} | \leqslant 15$ does not affect the region shown. Subtracting the background to obtain figure 10a reduces the peak height by a factor of almost 12, and reduces the width of the ridge from about $\pm$ 10 to $\pm$ 3. If we choose $Nq_{12} = 39$, at the high end of the observed distribution of overlaps, the triangle statistics again form a ridge with steep sides at $\pm$ 3-4 bonds. The subtracted data shown in figures 10b and 10c has

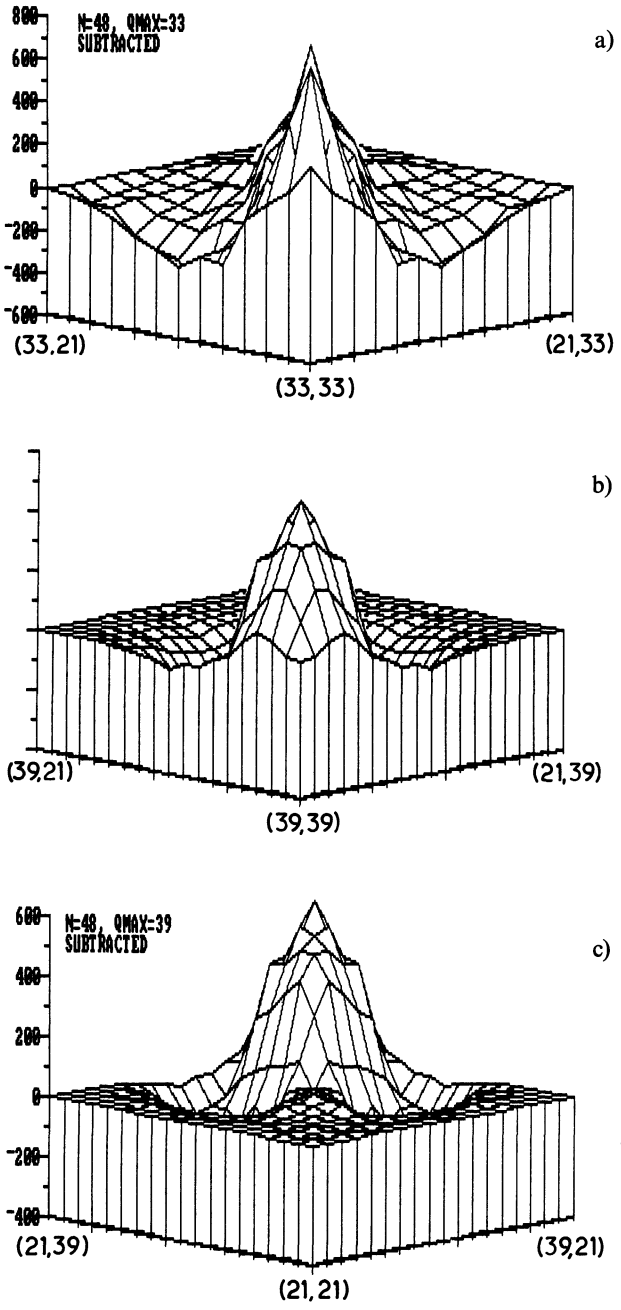


Fig. 10. — Subtracted statistics of all triangles formed among 3-opt tours of $N = 48$ cities with $Nq_{max} = 33$ (a) and 39 (b, c). Vertical scales are arbitrary.

a peak height which represents about half of the uncorrected data. Both front and back views (from $Nq_{13} = Nq_{23} = 39$ looking to small q, and from $Nq_{13} = Nq_{23} = 21$ looking to large q) are provided, since the negative density due to subtracting the background is largest in the middle of the surface plotted. The correlation that emerges in these plots is sensitive evidence for an ultrametric structure, and may be used for quantitative comparison with results in other models.

By integrating over the length of one side, q_{ik}, we obtain $P(q_{ij}, q_{jk}) - P(q_{ij})\,P(q_{jk})$ and collapse all our data onto a single plot for each value of N, eliminating the distraction of the triangle inequality. In figures 11a-f are displayed the results of this compression of the data for $N = 12$, 24 and 48. Both front views (down the $q_{ij} = q_{jk}$ axis) and side views (both q_{ij} and q_{jk} increase to the right) are shown. The distributions in figure 11 have a characteristic two-peaked form for all three problem sizes, and stay roughly constant in width as N increases. The upper peak has a greater width transverse to the ridge line than the lower peak, but this difference becomes less prominent as N increases.

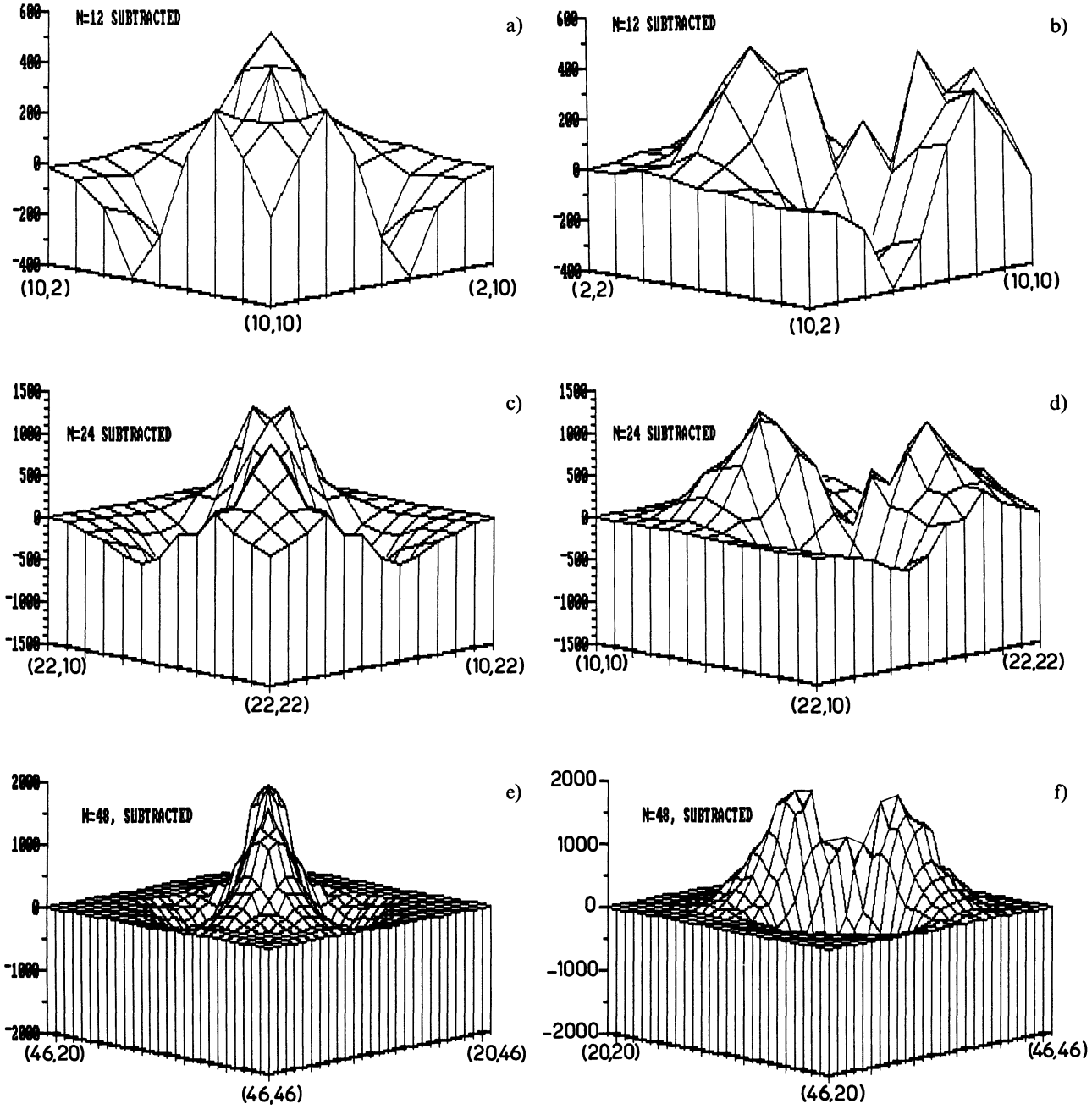


Fig. 11. — Subtracted statistics of pairs of bonds, for $N = 12$(a, b), $N = 24$(c, d), and $N = 48$(e, f). Figures 11a, c and e show the data looking along the $q_{ij} = q_{ik}$ axis from the large q end. Figures 11b, d and f show the same cases from the side. Vertical scales are arbitrary.

If instead of integrating over q_{12}, the strongest overlap, we integrate out both q_{13} and q_{23}, the function remaining (the fraction of all triangles with a given shortest side) can be directly compared with formulae derived for the S.K. model. The comparison is simplest if we consider the integrals of this function. The integral $\Sigma_<(q)$, which gives the fraction of triangles with all overlaps less than q, is plotted for an $N = 48$ instance in figure 12. For comparison we have plotted $x^3(q)$, the result expected if the overlaps are uncorrelated, and $x^2(q)$, the predicted value of $\Sigma_<(q)$ in the S.K. model, using $x(q) \equiv \int_0^q P(q)\, dq$ for the particular instance. The TSP results are closer to the uncorrelated formula than they are to the S.K. form, which is discussed in more detail in the following section. However, the difference between our result for $\Sigma_<(q)$ and the uncorrelated result is significant, and is seen for all three values of N. The S.K. result for $\Sigma_<$ depends sensitively upon a free parameter of the ultrametric statistics (relative weight of isosceles and equilateral triangles), so the comparison in figure 12 is not a clear-cut test for the presence or absence of such correlations.

Since the width of the various distributions discussed above is constant or very slowly increasing as N increases, it appears that the triangle and bond pair statistics will exhibit delta function contributions transverse to the $q_{ij} = q_{jk}$ axis in the limit $N \to \infty$. We offer no speculations on the finite size dependence in these results, and there is as yet no theory of finite size effects in the spin glass phase of the S.K. model.

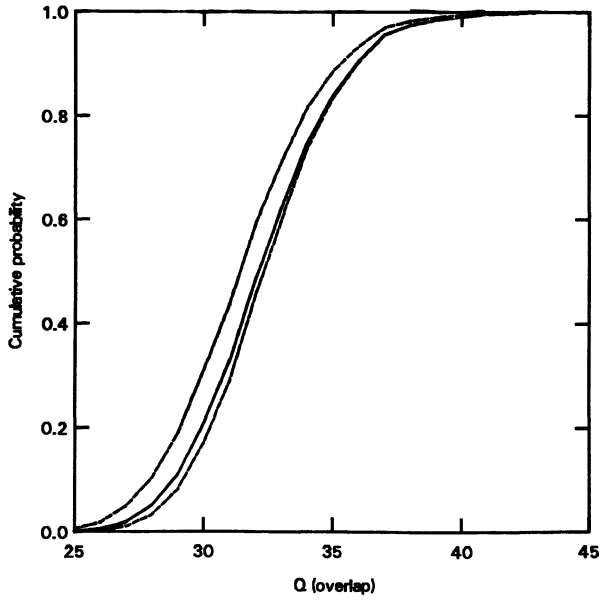


Fig. 12. — $\Sigma_<(q)$ defined as in the S.K. model, but for a random distance TSP with $N = 48$. For comparison, the S.K. result for infinite N, $x^2(q)$, is shown dashed, as is the uncorrelated limit $x^3(q)$.

For purposes of comparison, we calculated the 100-200 lowest energy spin configurations for S.K. models with 24 spins, using exact enumeration of all spin configurations. The data for $P(q_{ij}, q_{jk}) - P(q_{ij})\, P(q_{jk})$ for 4 instances are plotted in figure 13. While different in detail, the distribution has also two peaks, and roughly the same spread as our results for TSP.

Some of the general questions raised by this section are :

Is ultrametricity a generic property for random problems with large configuration spaces ? Could it be derived as a simple consequence of a law of large numbers, by maximizing some entropy of the configuration space landscape ?

To what extent are the various functions we used to characterize configuration space landscapes dependent on the choice of weights for the set of local minima which are retained in the analysis ?

3. The case of spin glasses.

The configuration space analysis for TSP in the two previous sections is natural and straight forward enough not to require exterior justifications. However it is historically true that some of the basic concepts are carry-overs from spin glass physics. And even if today the reference to spin glasses is not necessary it still provides an interesting and suggestive background for the general questions raised in section 2. As a consequence, in this section, we survey some of the basic relevant results from spin glass theory [22]. We shall only discuss long range Ising spin glasses.

The configuration space is an N-dimensional hypercube, with 2^N possible configurations. In the S.K. model [12], the Hamiltonian is

$$H = - \sum_{(ij)} J_{ij} S_i S_j \qquad (30)$$

where the interactions J_{ij} are independent random variables. For this model, a solution has been obtained in the large size limit (where the numbers of spins,

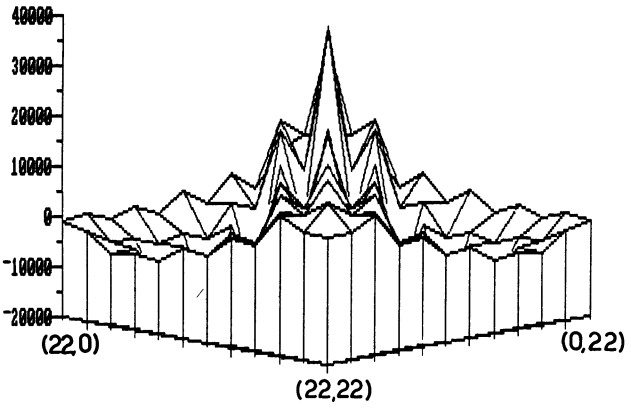


Fig. 13. — Overlap pair statistics, analogous to figure 11c but evaluated for an S.K. model with 24 spins. The vertical scale is arbitrary.

N, diverges). Some variants of the S.K. model (p-spin interactions, Potts or quadrupolar variables instead of spins) have also been considered.

The overlap between two spin configurations α and β is the scalar product in configuration space :

$$q^{\alpha\beta} = \frac{1}{N} \sum_i S_i^\alpha S_i^\beta \tag{31}$$

which takes values between -1 and $+1$. The fact that an overlap is an algebraic quantity, which can be positive and negative, is a noteworthy difference with our choice for the TSP overlap. Physically, the need for an algebraic overlap in the spin glass case comes from time-reversal invariance. To any spin configuration corresponds a time-reversed configuration with all spins flipped. In the presence of a magnetic field, time reversal invariance is broken, and it turns out that only positive overlaps remain of importance in the thermodynamic analysis. Thus the best analogy to TSP is a spin glass with an applied magnetic field.

For the S.K. model, there exists a sharp phase transition temperature T_g (in the large N limit) below which the system is not ergodic. Thus for $T < T_g$, the system gets trapped into one or another region of configuration space from which it cannot escape toward other regions with comparable energy ; this is the definition of ergodicity breaking. Configuration space is thus divided into valleys separated by barriers too high to be passed over. Of course, for a finite system, there is always the possibility to climb over barriers because they are finite. That is the reason why sharp phase transitions can only occur in the large size limit. Therefore the concept of valleys must be used with some care in the context of finite systems.

Within one valley s, one can define the average magnetization of spin S_i as being :

$$m_i^s = \langle S_i \rangle_s$$

and the overlap between two valleys s and s' is

$$q^{ss'} = \frac{1}{N} \sum_i m_i^s m_i^{s'} . \tag{32}$$

Now, one can define a valley overlap distribution function by

$$P(q) = \sum_{s,s'} W_s W_{s'} \delta(q - q^{ss'}) \tag{33}$$

where W_s is the Boltzmann weight of valley s

$$W_s = \frac{1}{Z} \sum_{\alpha \in s} e^{-E^\alpha/T}$$

$$\sum_s W_s = 1 . \tag{34}$$

Note that this Boltzmann weight of a valley is a single number which is influenced by the height, the size and more generally the shape of the valley. It does not exhaust all physical information one might wish to obtain on a valley.

The number of valleys and the shape of $P(q)$ vary with temperature and field in the spin glass phase. In the S.K. model, $P(q)$ contains a continuous component. But, there exists a simpler spin glass model, the random energy model, where $P(q)$ contains only δ-function peaks, as in a standard phase transition [24]. Formulae below pertain to both models, with the inclusion of the appropriate $P(q)$.

The ultrametricity property was discovered by computing $P(q_1, q_2, q_3)$ which gives the statistics of triangles formed by picking three valleys at random [10]. Indeed, it was found that :

$$P(q_1, q_2, q_3) = \sum_{s,s',s''} W_s W_{s'} W_{s''} \delta(q_1 - q^{ss'}) \delta(q_2 - q^{s's''}) \delta(q_3 - q^{ss''})$$

$$= 1/2 \, [P(q_1) P(q_2) \theta(q_1 - q_2) \delta(q_2 - q_3) + \text{permutations}]$$

$$+ 1/2 \, P(q_1) x(q_1) \delta(q_1 - q_2) \delta(q_2 - q_3) , \tag{35}$$

$$\text{where } x(q) = \int_0^q P(q) \, dq .$$

Formula (35) shows that the triangles are either isosceles with small basis, or equilateral, which is the defining property of a ultrametric space.

This ultrametricity property suggests that, as temperature is lowered, the ergodically separated valleys are generated by a branching process, with only one valley at $T > T_g$ (paramagnetic phase) and more and more valleys as T decreases (of course, the weights are temperature dependent). Conversely, assuming the existence of such an evolution tree and that the overlap between two valleys is determined by the distance to their closest common ancestor, the ultrametricity property follows.

There are various ways to prove that the weight and environment of a valley do vary considerably from one to the other. An exact expression [10] has been given for the fluctuations of $P(q)$ over s :

$$P^s(q) = \sum_{s'} W_{s'} \delta(q - q^{ss'}) . \tag{36}$$

This imbalance between valleys is actually hidden in the coefficients of formula (35). If all valleys were equivalent (same weight W_s and same environment as described by the function $P^s(q)$), the integrated function :

$$P(q_1, q_2) = \int_0^1 P(q_1, q_2, q_3)\, dq_3 \tag{37}$$

would factorize into $P(q_1)\, P(q_2)$. That is not the case for spin glass since, using (35), one finds :

$$P(q_1, q_2) = 1/2\, P(q_1)\, P(q_2) + 1/2\, P(q_1)\, \delta(q_1 - q_2)\,. \tag{38}$$

The function $C(q_1, q_2) = P(q_1, q_2) - P(q_1)\, P(q_2)$ therefore deserves consideration. A word on normalization is in order. The function $P(q)$ should include self-overlaps in order to be properly normalized to 1. $P(q_1, q_2, q_3)$ should also include all triangles, including triangles in which the same valley occurs two or three times, in order to be conveniently normalized. With these precautions :

$$\int_0^1 C(q_1, q_2)\, dq_2 = 0\,. \tag{39}$$

In spin glasses, as in TSP, the function $C(q_1, q_2)$ exhibits positive values on the diagonal surrounded by negative values. The formula for spin glasses (38) does not distinguish between weight and environment fluctuations. In the numerical analysis of section 2, for the TSP, equal weight was taken for locally stable configurations. In order to compare comparable things, we have used similar weighting in figure 13 which is to be compared with figures 11. In both cases, the source of imbalance comes from environment fluctuations.

Other quantities, interesting and simple to compute, are $\Sigma_<(q)$ and $\Sigma_>(q)$ defined by :

$$\Sigma_<(q) = \int_0^q dq_1 \int_0^q dq_2 \int_0^q dq_3\, P(q_1, q_2, q_3)\,, \tag{40}$$

$$\Sigma_>(q) = \int_q^1 dq_1 \int_q^1 dq_2 \int_q^1 dq_3\, P(q_1, q_2, q_3)\,, \tag{41}$$

which count the proportion of triangles with all three sides larger or smaller than q. From formula (35), one gets for S.K. spin glasses :

$$\Sigma_<(q) = x^2(q) \tag{42}$$

$$\Sigma_>(q) = [1 - x(q)]\left[1 - \frac{x(q)}{2}\right] \tag{43}$$

to be compared with the result obtained assuming no correlation between the three triangle side lengths.

$$\Sigma^u_<(q) = x^3(q) \tag{44}$$

$$\Sigma^u_>(q) = [1 - x(q)]^3\,. \tag{45}$$

We have been considering increasingly global properties of the valley distribution. Finally, the most global property is the statistical entropy of the distribution of valleys, which is defined as usual by

$$K = -\sum_s W_s \ln W_s\,. \tag{46}$$

In order to distinguish it from the total entropy, which contains contributions coming from the intra-valley entropy, one sometimes calls K the complexity of the valley distribution [22]. In the S.K. model, the complexity is zero in the paramagnetic phase since there is only one valley. It rises below T_g but it is a quantity of order O(1) and not of order O(N), due to the Boltzmann weighting. The interpretation is that there are some dominant valleys of finite weight among the O(exp N) valleys. In zero field, the complexity jumps from 0 to ln2 at T_g, and then rises smoothly as the temperature decreases. The jump at T_g is due to the spontaneous breaking of time reversal invariance (ferromagnetic-like transition), which somewhat accidentally occurs on top of the spin glass transition. This peculiarity of spin glasses in zero field is not relevant for TSP which, as discussed previously, is best compared with a spin glass in the presence of a magnetic field.

As mentioned earlier, the notion of a valley is somewhat ill-defined at finite temperatures for a finite sample. Still, some of the previous quantities can be unambiguously defined. For instance, $P(q)$ can be defined from

$$P(q) = \sum_{\alpha,\beta} W_\alpha\, W_\beta\, \delta(q - q^{\alpha\beta}) \tag{47}$$

where W_α is the Boltzmann weight of configuration α. Numerical estimates of $P(q)$ have been obtained with the Monte Carlo algorithm for long range [25] and finite range [26] spin glass.

Another approach, closer to our TSP analysis, consists in focussing on locally stable configurations at zero temperature, the so called TAP solutions [22] in the context of spin glasses. Equal weight can be given to all TAP solutions below some energy cutoff, and zero weight above the cutoff. There are no analytical results for such a choice of weights, but these statistics are obviously of interest for the analysis of finite systems. Alternately the finite temperature thermodynamic properties can also be reconstructed from the numerical analysis of TAP solutions, with some approximations. The results appear to be in good agreement with theoretical predictions [27]. All this suggests that a detailed analysis of the distribution of TAP solutions in configuration space is not only useful to check the thermodynamic predictions and the

validity of the existing theory but would also provide additional information of interest. One would like to know, for instance, the statistics of the attractor basins of locally stable solutions. The basin size is a useful weight for the models of memory proposed by Hopfield [28]. The lowest valley, which will eventually dominate with Boltzmann weighting, need not have the largest attractor basin [31]. So different weightings may provide complementary information, as discussed in the next section.

We end this section with two comments on the connections between spin glass and algorithmic complexity.

i) It has been proved that the spin glass problem defined as an optimization problem (i.e., find the ground state) is *NP*-complete in dimensions larger than 2 [29, 30]. Though it is polynomial in dimension 2, it becomes *NP*-complete for two coupled planes [29]. It is tempting to link *NP*-completeness with the occurrence of a sharp phase transition at finite temperatures in the thermodynamic (large size) limit, but growing evidence suggests to resist this temptation [32].

It does seem that *NP*-completeness is generally associated with some sort of ergodicity breaking, with a freezing transition which may be sharp or smooth, and when smooth with a zero temperature phase transition, as in the case of the TSP. However, these characteristics are not exclusive to *NP*-complete problems, since the 2 D spin glass problem, which is in *P*, possesses them too. All this suggests that a more refined categorization of *NP*-complete problems is needed.

ii) Some classic optimization problems, such as min-cut partitioning, are now recognized as being simply expressible as spin glass problems [3].

4. Perspective and questions.

In the previous sections, a number of tools for the analysis of configuration space landscapes have been discussed. Some arose from the study of spin glasses (ultrametricity tests, triangle statistics...). Our approach to the TSP has introduced different (non-Boltzmann) weightings, which are also of physical interest. Spin glass theory led us to put emphasis on the overlap of two configurations. In the TSP context, where the overlap of two configurations is the number of common bonds, it appears natural to consider also higher correlations [15], such as the number of bonds common to three tours, etc. It shows that the battery of tools relevant for the study of configuration space landscapes is likely to expand as more optimization problems are explored.

While the means of analysis are sharpened, it is worth bringing this knowledge to bear on the modification of configuration space landscapes under changes of the parameters defining an instance. There are several reasons to be interested in this problem of landscape « gardening ».

Firstly, the lack of self-averaging in the mean field theory of spin glass suggests that there may exist a considerable amount of plasticity in configuration space landscapes, making them sensitively dependent on rather minor modifications.

Secondly, configuration space gardening is related to the notions of learning and unlearning for artificial or animal memories. Hopfield's model of a content-addressable memory is essentially an Ising spin model, with spins representing neurons and interactions representing synapses. Under suitable symmetric assumptions $J_{ij} = J_{ji}$, the neuron dynamics is a gradient dynamics, running downhill in configuration space. However, the outcome of Hopfield's learning and unlearning process is a configuration space evenly divided into valleys of similar size. In a phase transition language, this would correspond to an evolution tree with one branching temperature, namely only one relevant temperature scale. However, a hierarchical structure of configuration space is physiologically more appealing, because it allows for a safer and more resourceful classification of memories. Since a purely random distribution of connection strengths (spin glass limit) possesses this property, it appears that better advantage should be taken of randomness. Thus it is tempting to explore ways of pruning the configuration space of a spin glass from its excessive number of valleys while keeping its hierarchical properties.

An apparent objection against this approach comes from the theoretical suggestion that there is a small number of dominant valleys (obsessions) in long range spin glasses. It should be recognized however that this dominance is due to the Boltzmann weighting which is used in statistical physics. In the context of memories, it is the attractor basin size which is a proper measure of the weight of a valley. Present numerical analysis indicates that with such weighting the obsession problem disappears.

As a conclusion, we believe that the main contribution of this paper consists in the new questions it raises.

Acknowledgments.

Fruitful discussions with P. W. Anderson, J. P. Benzécri, M. Mézard, R. Rammal, N. Sourlas, J. Vannimenus, M. Virasoro, and A. P. Young are gratefully acknowledged.

References

[1] PAPADIMITRIOU, C. H., STEIGLITZ, K., *Combinatorial Optimization* (Prentice Hall) 1982.

[2] KIRKPATRICK, S., *Lectures Notes in Physics,* vol. 149 (Springer) 1981, p. 280 ;
KIRKPATRICK, S., GELATT Jr., C. D., VECCHI, M. P., *Science* **220** (1983) 671 (this last paper is referred to as KGV in the text).

[3] KIRKPATRICK, S., *J. Stat. Phys.* **34** (1984) 975.

[4] ČERNÝ, V., preprint (1982), to appear in *J. Optimization Theory Appl.* (1985).

[5] BONOMI, E., LUTTON, J. L., *SIAM Rev.* **26** (1984) 551.

[6] LUNDY, M., MEES, A., preprint (1984).

[7] SKISCIM, C. C., GOLDEN, B. L., preprint (1983).

[8] JOHNSON, D. S., private communication.

[9] PARISI, G., *Phys. Rev. Lett.* **50** (1983) 1946.

[10] MÉZARD, M., PARISI, G., SOURLAS, N., TOULOUSE, G., VIRASORO, M., *Phys. Rev. Lett.* **52** (1984) 1146 ; *J. Physique* **45** (1984) 843.

[11] LIN, S., *Bell Syst. Tech. J.* **44** (1965) 2245.

[12] BEARDWOOD, J., HALTON, J. H., HAMMERSLEY, J. M., *Proc. Cambridge Philos. Soc.* **55** (1959) 299.

[13] SHERRINGTON, D., KIRKPATRICK, S., *Phys. Rev. Lett.* **35** (1975) 1792.

[14] VANNIMENUS, J., MÉZARD, M., *J. Physique Lett.* **45** (1984) L-1145.

[15] GROSS, D., MÉZARD, M., private communication.

[16] TROTTER, H. F., Algorithm 115, *Commun. ACM* **5** (1963) 434.

[17] DE DOMINICIS, C., GABAY, M., GAREL, T., ORLAND, H., *J. Physique* **41** (1980) 923.

[18] BRAY, A. J., MOORE, M. A., *J. Phys.* **C 13** (1980) L469.

[19] BENZÉCRI, J. P., *La Taxonomie* (Dunod) 1984.

[20] JARDINE, N., SIBSON, R., *Mathematical Taxonomy* (Wiley) 1971.

[21] SOKAL, R. R., SNEATH, P. H. A., *Principles of Numerical Taxonomy* (Freeman) 1963.

[22] Heidelberg Colloquium on Spin Glasses, *Lecture Notes in Physics*, Vol. 192 (Springer) 1983.

[23] YOUNG, A. P., BRAY, A. J., MOORE, M. A., *J. Phys.* **C 17** (1984) L143.

[24] GROSS, D., MÉZARD, M., *Nucl. Phys.* **B 240** [FS12] (1984) 431.

[25] YOUNG, A. P., *Phys. Rev. Lett.* **49** (1983) 1206.

[26] SOURLAS, N., *J. Physique Lett.* **45** (1984) L-969.

[27] PARGA, N., PARISI, G., VIRASORO, M., *J. Physique Lett.* **45** (1984) L-1063.

[28] HOPFIELD, J. J., *Proc. Natl. Acad. Sci. USA* **79** (1982) 2254.

[29] BARAHONA, F., MAYNARD, R., RAMMAL, R., UHRY, J. P., *J. Phys. A* **15** (1982) 673.

[30] BACHAS, C., *J. Phys.* **A 17** (1984) L 709.

[31] VIRASORO, M., private communication.

[32] ANGLES D'AURIAC, J. C., PREISSMANN, M., RAMMAL, R., *J. Physique Lett.* **46** (1985) L-173.