



HAL
open science

On the degree of ultrametricity

R. Rammal, Jean-Christian Anglès d'Auriac, Benoît Douçot

► **To cite this version:**

R. Rammal, Jean-Christian Anglès d'Auriac, Benoît Douçot. On the degree of ultrametricity. Journal de Physique Lettres, 1985, 46 (20), pp.945-952. 10.1051/jphyslet:019850046020094500 . jpa-00232922

HAL Id: jpa-00232922

<https://hal.science/jpa-00232922>

Submitted on 4 Feb 2008

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

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

LE JOURNAL DE PHYSIQUE - LETTRES

J. Physique Lett. **46** (1985) L-945-L-952

15 OCTOBRE 1985, PAGE L-945

Classification

Physics Abstracts

02.70 — 05.20 — 05.50

On the degree of ultrametricity

R. Rammal, J. C. Angles d'Auriac and B. Doucot

Centre de Recherches sur les Très Basses Températures, CNRS,
B.P. 166 X, 38042 Grenoble Cedex, France

(Reçu le 17 juillet 1985, accepté le 30 août 1985)

Résumé. — La notion de degré d'ultramétrie \mathcal{D} d'un espace métrique donné est introduite à partir de l'ultramétrie sous-dominante. On décrit une procédure simple et efficace pour le calcul de \mathcal{D} . On montre que \mathcal{D} fournit une mesure quantitative simple de la déviation par rapport à l'ultramétrie exacte. Cette notion est illustrée pour des exemples explicites et nous suggérons son intérêt dans les modèles de mécanique statistique ainsi que les problèmes d'optimisation combinatoire.

Abstract. — Using the notion of the subdominant ultrametric, the degree of ultrametricity \mathcal{D} of a given metric space (e.g. phase space) is introduced. A simple and efficient method for the calculation of \mathcal{D} is outlined. \mathcal{D} is shown to provide a simple quantitative measure of the deviation from exact ultrametricity. Explicit examples are used to illustrate this notion which is argued to be of some interest in statistical-mechanical models and combinatorial optimization problems.

1. Introduction.

During the past few years, a considerable amount of work has been done in the study of spin glass models. In particular, it has been suggested [1] that configuration space of pure states is endowed with a non-trivial ultrametric structure [2]. This suggestion has been checked numerically by various authors [3] for different spin glass models. The same kind of configuration space analysis has also been carried out for some combinatorial optimization problems [4]. In both cases, the direct method of the statistics of triangles has been used in order to check the ultrametricity of configuration spaces. In finite spin models for instance, a finite set of representative configurations is used to perform such an analysis usually. A further step in that direction would be a quantitative measure of the degree of ultrametricity of a given configuration space. The control of finite size effects is just an example of a practical use of such notion. Actually, the ultrametric inequality [5] is so stringent (all triangles are either isosceles with a small base or equilateral) that, the modification of one spin configuration only may result in the loss of this property.

This may occur for instance in spin models because of thermal fluctuations which lead to fluctuating distances. Whereas at zero temperature the search for ultrametricity in the whole set of ground states is actually altered by finite size effects.

In this paper, we present a new method for the measurement of the degree of ultrametricity of a given configuration space. This method is a very general one and can be used in various contexts. The basic ingredient is the notion of the subdominant ultrametric. This concept, widely used in Taxonomy [6] is recalled in the next section. In order to be self-contained, the main properties of subdominant ultrametries will be summarized in that section. A simple and efficient method for the calculation of the degree of ultrametricity, for a given metric space, is outlined. Illustrative examples are the object of the remaining part of the paper.

2. The subdominant ultrametric.

First we define the problem and recall some basic results. Let us consider a finite metric space (Ω, d) , where the distance between $x \in \Omega$ and $y \in \Omega$ is denoted by $d(x, y)$. In order to measure the degree of ultrametricity of (Ω, d) one asks for an ultrametric denoted (UM hereafter) d' , defined on Ω such as d' is the closest UM to d . This naturally defines an optimization problem on the set U of ultrametries defined on Ω . The pertinent formulation of such a problem suffers actually from two intrinsic difficulties. The first one is the choice of the measure of proximity of two metrics d and d' on Ω . The second one is the problem of unicity of d' . An elegant and simple solution, but still a partial one, is provided by the notion of the subdominant ultrametric.

Formally, the subdominant UM associated with d can be defined as follows. Note first that the whole set of metrics, defined on Ω , can actually be ordered in a natural way, as :

$$d' \leq d \Leftrightarrow d'(x, y) \leq d(x, y) \quad \text{for all } (x, y) \in \Omega \times \Omega.$$

Consider now the set $U^< = \{d' \in U \mid d' \leq d\}$ of the ultrametries on Ω which are smaller than d . The maximal element $d^<$ of $U^<$ exists and can be shown to be unique [6]. The ultrametric $d^<$ so defined is called the subdominant UM associated with d . Note that the restriction $d' \leq d$ in the search for $d^<$ is a very important condition, insuring the unicity of $d^<$. In this respect, it is important to notice that the minimal element of $U^> = \{d' \in U \mid d' \geq d\}$ is not uniquely defined. Before describing a simple method for the construction of $d^<$, let us mention two important properties of $d^<$.

i) *Optimality.* — Assume that the proximity Δ of two metrics d and d' is measured by the Minkowski distance :

$$\Delta_\infty(d, d') = \text{Max}_{x, y \in \Omega} |d(x, y) - d'(x, y)|$$

or

$$\Delta_\alpha(d, d') = \left[\sum_{x, y \in \Omega} |d(x, y) - d'(x, y)|^\alpha \right]^{1/\alpha}, \quad \alpha > 0.$$

Then, among all the elements of $U^<$, $d^<$ is the only ultrametric for which $\Delta_\infty(d, d')$ and more generally $\Delta_\alpha(d, d')$ assume their minimal values [7].

ii) *Stability.* — If d is transformed into λd ($\lambda > 0$), i.e. : $d(x, y) \rightarrow \lambda d(x, y)$, $x \in \Omega$, $y \in \Omega$, then $\lambda d^<$ is the subdominant ultrametric associated with λd . More generally, this stability property remains true when d is transformed into $\varphi(d)$ where φ denotes a positive non-decreasing function such that $\varphi(a + b) \leq \varphi(a) + \varphi(b)$, $a \geq 0$, $b \geq 0$.

Given the metric d on Ω , there are different methods [8] for the construction of $d^<$. In what follows, we shall describe a simple procedure giving $d^<$ in a very transparent way : the minimal

spanning tree construction method. We shall outline briefly the main step of this construction using a graph theoretic language. To a given metric space (Ω, d) corresponds a simple, undirected graph whose vertices are the elements of Ω . A length $d(x, y)$ is assigned to each edge (x, y) . The basic step for obtaining $d^<$ is the construction of a minimal spanning tree (MST) on the connected graph structure so obtained. Recall [9] that MST refers to a tree A , having the same vertices as Ω , but of minimal total length. Note that A may not be uniquely defined and more than one MST can be constructed on Ω . However, despite this possible non unicity of A , $d^<$ as obtained below, is unique and does not depend on the particular choice of A . Having a MST on Ω , the distance $d^<(x, y)$ between two elements x and y in Ω is given by :

$$d^<(x, y) = \text{Max} \{ d(w_i, w_{i+1}), 1 \leq i \leq n - 1 \}$$

where $(w_1 = x, w_2), (w_2, w_3), \dots, (w_{n-1}, w_n = y)$ are the $n - 1$ successive edges of the unique chain between x and y on the tree. This defines precisely the subdominant ultrametric $d^<$.

A very simple procedure for constructing a MST on Ω is provided by Kruskal's algorithm [9] which can be summarized as follows. Start with $A = \emptyset$ (the empty set) and order the set of distances $d(w_i, w_j)$ of Ω in increasing order. Each distance $d(w_i, w_j)$ is associated with an edge (w_i, w_j) on the graph of Ω . Put in A the first edge of the list. Next edges in the list are then added successively to A , such as to avoid the formation of cycles with edges already present in A . This procedure will stop after $|\Omega| - 1$ steps, giving a MST on Ω and then $d^<$. Schematic examples are shown in figures 1 and 2.

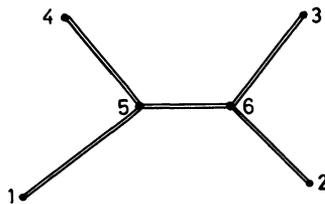


Fig. 1. — A schematic example of $\Omega = \{ 1, 2, 3, 4, 5, 6 \}$ showing a minimal spanning tree (MST), represented by double lines, used in the construction of the subdominant ultrametric $d^<$. For instance, $d^<(1, 2) = \text{Max} \{ d(1,5), d(5,6), d(6,2) \}$, $d^<(2,3) = \text{Max} \{ d(2,6), d(6,3) \}$.

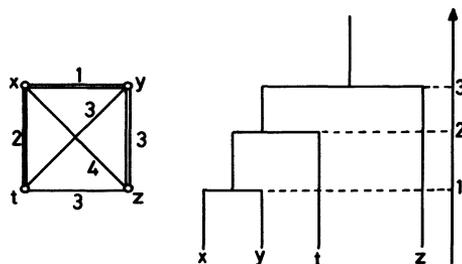


Fig. 2. — Example of a metric space (Ω, d) , $\Omega = \{ x, y, z, t \}$ with the associated hierarchy defined by $d^<$. Double lines denote the edges of a MST on Ω . In this example, the degree of ultrametricity \mathfrak{D} of (Ω, d) is given by $\mathfrak{D} = 1 - \frac{14}{16} = \frac{1}{8}$.

3. Degree of ultrametricity. Example.

How can we define and measure a degree of ultrametricity for a metric space (Ω, d) which is not exactly ultrametric ? What is the « minimal distortion » of d in order to become an ultrametric ? One possible answer to these questions is actually provided by the direct comparison of d with $d^<$. A measure of the proximity of d and $d^<$ will be given for instance by one of the Minkowski's distances Δ_x defined above. For the sake of simplicity, the following expression for the degree of ultrametricity \mathfrak{D} , or relative distortion, will be used :

$$\mathfrak{D} = \frac{\sum_{x,y \in \Omega} (d(x,y) - d^<(x,y))}{\sum_{x,y \in \Omega} d(x,y)}. \quad (1)$$

Here d is the input metric on Ω and $d^<$ is the ultrametric characterizing the output hierarchy on Ω . In general, \mathfrak{D} is such that $0 \leq \mathfrak{D} \leq 1$, and vanishes if d is initially an ultrametric ($d^< = d$). \mathfrak{D} as defined by equation (1) provides actually a quantitative measure of the degree of ultrametricity of (Ω, d) . Small values of \mathfrak{D} indicate that d is not very far from being ultrametric, this being realized through a slight distortion of d . Whereas values of \mathfrak{D} close to one tell us that large distortions of d are actually needed in order to transform d into an ultrametric.

EXAMPLE. — A simple class of examples, where the construction of a MST is trivial, is provided by $\Omega = \{x_1, x_2, \dots, x_n\}$ where x_i denotes n points on the real line. The distance $d(x_i, x_j) = |x_i - x_j|$ is the Euclidean distance. In the particular case, where $x_i = i$ ($1 \leq i \leq n$), the subdominant ultrametric $d^<$ reduces to the trivial ultrametric $d^<(x_i, x_j) = 1$: all triangles are equilateral. For this example, $\mathfrak{D} = 1 - 3/(n+1) \sim 1$ at large n , in perfect agreement with the intuitive expectation : Euclidean spaces are far from being ultrametric spaces. This conclusion is not modified when $\Omega = \{x_1, x_2, \dots, x_{n+1}\}$ corresponds to $(n+1)$ randomly chosen points on the real line. Indeed, let us denote by $X_1 \leq X_2 \leq \dots \leq X_{n+1}$ the corresponding ordered sequence and let $L_k = X_{k+1} - X_k$, $k = 1, 2, \dots, n$, be the length of the interval $[X_k, X_{k+1}]$. Then a simple calculation leads to the following expression for \mathfrak{D} : $\mathfrak{D} = 1 - S'/S$, where $S = \sum_{k=1}^n k(n+1-k) L_k$

and $S' = \sum_{k=1}^n \sum_{l=k}^n \text{Max}(L_k, \dots, L_l)$ respectively. Assuming that the L_k 's are n independent random variables having a probability density $p(L) = dF/dL$, then the average deviation from d to $d^<$ takes the following form : $\langle S - S' \rangle = \sum_k (n+1-k) (k \langle L \rangle - \mu_k)$. Here $\langle L \rangle = \int L dF(L)$ and μ_k is the average value of $\text{Max}(L_1, \dots, L_k)$. For non-singular $p(L)$, $\langle S \rangle \sim n^3$ ($n \gg 1$) and $\langle S' \rangle$ can also be calculated : $\langle S' \rangle \sim n^2$ for uniform $p(L)$, $\langle S' \rangle \sim n^2 \ln(n)$ for exponential $p(L)$, etc. In all these cases, as well as for other variants (e.g. different probability distribution for $\{x_i\}$) of this example, one has $\mathfrak{D} \sim 1$ at large n , as anticipated above.

4. Explicit example.

A more suggestive example is given by the set of B -bit words. As a space Ω we shall consider a subset consisting of M words, randomly taken among the whole set of $V \equiv 2^B$ possible B -bit words. The elements of Ω can be viewed as spin configurations of a system consisting of B Ising spins. The set of 2^B words represents the vertices of the elementary Euclidean hypercube in B dimensions. The distance (Hamming) between two words will be defined by : $d(x, y) = \sum_{i=1}^B |x_i - y_i|$, for $x = (x_1, x_2, \dots, x_B)$, $y = (y_1, y_2, \dots, y_B)$ in Ω , $x_i \in \{0, 1\}$, $y_i \in \{0, 1\}$. The relevant questions are the following : How does the degree of ultrametricity $\mathfrak{D}_B(x)$ behave as a function of B and the filling factor $x = M/2^B$? What is the asymptotic behaviour ($B \gg 1$) of $\mathfrak{D}_B(x)$?

4.1 GENERAL FEATURES. — The following two limits are easy to understand. The first one corresponds to $x = 1$, i.e. $M = 2^B$. In this limit, $d^<$ is trivially given by $d^<(x, y) = 1$ and $\mathcal{D}_B(1) = 1 - 2/B$. The second limit corresponds to a small number of words : $x \ll 1$. For instance, at $M = 3$, the probability distribution of the triangle sides is binomial, with an average $d(x, y) \sim B/2$ and a variance $\sigma \sim B$. This leads to $\mathcal{D}_B(x \ll 1) \sim B^{-1/2}$ showing that, in spaces with large B , almost all triangles are equilateral and ultrametricity is obtained in a rather natural way.

For a fixed B , the function $\mathcal{D}_B(x)$ is expected to increase uniformly with x , in order to interpolate between these two limits. The limiting behaviour at $B = \infty$ is a trivial one, with $\mathcal{D}_\infty(x) = 0$ for $x = 0$ and $\mathcal{D}_\infty(x) = 1$ for $x \neq 0$. This is easily understood because, for a given $x = M/2^B$, $\mathcal{D}_B(x)$ must be an increasing function of B . In fact assume that $B \rightarrow B + 1$ and M is doubled. The new $2M$ words, having each one more bit (the last one) can be obtained from the previous M ones, by adding 0 or 1 as a new bit. This is not so restrictive an assumption, and the new hierarchy can be deduced from the one previously known. Indeed, one can show :

$$\mathcal{D}_{B+1}(x) = 1 - [4 \sum d^<(i, j) + M] / [4 \sum d(i, j) + M^2]$$

which implies $\mathcal{D}_{B+1}(x) > \mathcal{D}_B(x)$, at least for large M . Here the sums are taken over the set of couples (i, j) in Ω .

4.2 NUMERICAL STUDY. — The degree $\mathcal{D}_B(x)$ has been calculated numerically, in the whole interval $0 \leq x \leq 1$, for different values of $B : 8 \leq B \leq 160$. Typical results are shown in figures 3 and 4. Instead of a smoothly behaved function, $\mathcal{D}_B(x)$ actually exhibits a staircase behaviour. Already for small B ($B \sim 10$), $\mathcal{D}_B(x)$ is characterized by jumps between « quantized » values : $1 - 2k/B$ ($k \geq 1$), the stairs are stable (small fluctuations from sample to sample) and the jumps become sharper and sharper when B increases and appear at very precise values of x . For small

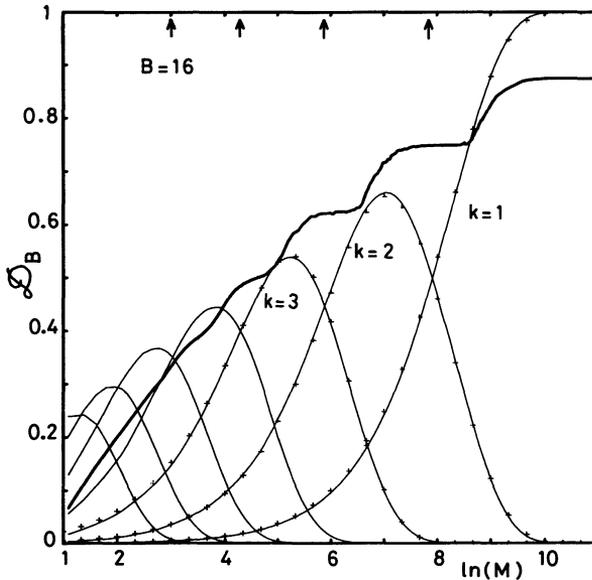


Fig. 3. — Step-like behaviour (heavy solid line) of the degree of ultrametricity $\mathcal{D}_B(x)$, calculated for 500 samples of M words ($B = 16$). The probability distribution P_k for the nearest-neighbour distances is shown for different values of k . Solid lines represent the calculated distributions whereas points (+) reproduce the expression given in the text (Eq. (2)). Arrows indicate the centres of stairs defined by $P_k = P_{k-1}$.

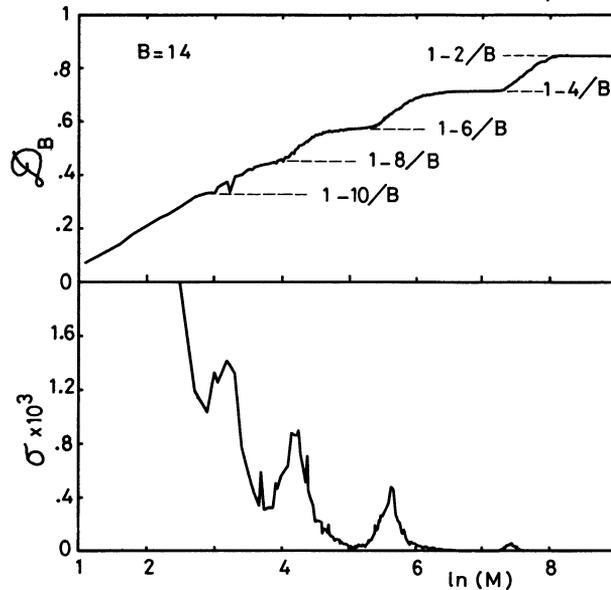


Fig. 4. — The same $\mathcal{D}_B(x)$ as in figure 3, for $B = 14$. The lowest part shows the variance σ of the same quantity calculated for 500 samples.

values of x , $\mathcal{D}_B(x)$ becomes smooth, due to finite size effects and for large x , $\mathcal{D}_B(x)$ ends with a stair corresponding to the expected value $\mathcal{D}_B(x) = 1 - 2/B$. Each jump between two stairs is actually accompanied with a large increase of the magnitude of sample to sample fluctuations. This is clearly shown in figure 4, where the calculated variance σ is shown together with $\mathcal{D}_B(x)$ for $B = 14$.

This behaviour is actually obtained for larger values of B , where an increasing number of robust stairs and sharp jumps appear. This picture becomes clearer for increasing B with the appearance of successive plateaus. These jumps are reminiscent of instability phenomena and, as will be shown below, the robust stairs can be viewed as coexistence regions.

4.3 GEOMETRICAL PICTURE. — The obtained behaviour of $\mathcal{D}_B(x)$ can be understood by keeping in mind the MST construction method and the sphere coverage problem [10]. Using the representation of Ω as a part of the B -dimensional hypercube, it is useful to consider each element of Ω as the centre of a k -sphere, having a radius equal to an integer k ($1 \leq k \leq B/2$). This defines a neighbour for each point in Ω , containing $w_k = \sum_{i=0}^k \binom{B}{i}$ points of the hypercube. The coverage problem of the hypercube is the following. For M randomly chosen centres and a given value of k , find the probability p_M for a complete coverage with the Mk -spheres. Clearly, $p_M \leq 1 - (1 - w_k/2^B)^M$, and for large M but small $w_k/2^B$, $p_M \lesssim 1 - \exp(-x \cdot w_k)$. As will be shown below, the particular values of x , $x_k \sim 1/w_k$ are called to play a key role in the behaviour of $\mathcal{D}_B(x)$.

Now consider the probability distribution of the nearest-neighbour distances in Ω , as a function of M and B . In fact, the structure exhibited by $\mathcal{D}_B(x)$ turns out to be intimately related to that probability distribution. Given a point in Ω , the probability that all the $M - 1$ other points are at a distance strictly larger than k is given by: $\pi_k \sim (1 - w_k/2^B)^{M-1}$. This expression becomes accurate when the overlap between the k -spheres is negligible. Therefore, the probability distri-

bution for the nearest-neighbour distances is given by

$$P_k \equiv \pi_{k-1} - \pi_k = (1 - w_{k-1}/2^B)^{M-1} - (1 - w_k/2^B)^{M-1}. \tag{2}$$

This expression is shown in figure 3, for different values of k ($k \geq 1$). The numerical results, for different values of k , M and B follow equation (2) and small deviations are obtained for large k and small x .

The first stair at $1 - 2/B$ actually corresponds to a close packing configuration, where the hypercube is covered, with a probability ~ 1 , by 1-spheres. Indeed, a simple calculation, giving the corresponding MST, leads to $\mathcal{D}_B(x) = 1 - 2/B$ in that region. This stair ends up at a value of M (see below) given by $M/2^B \sim 1/w_1$, i.e. $x \sim 1/(B + 1)$.

For each value $k \geq 2$, P_k assumes its maximal value at $x = \alpha_k$, where almost all nearest-neighbour distances are equal and the hypercube is almost surely covered by k -spheres. These values of the filling factor are associated with the unstability of regular hierarchies and give the positions of the jumps. Indeed, adding (or removing) few points in Ω results in a strong modification of the hierarchy, making $\mathcal{D}_B(x)$ have a sharp jump. Large fluctuations occur in the neighbourhood of $x = \alpha_k$, because of this unstability in the hierarchy associated with subdominant ultrametrics.

The centres of stairs can also be calculated from the condition : $P_k = P_{k-1}$. Indeed, for such values x_k defined by this condition, the two distances k and $k - 1$ are equiprobable and are the most probable ones. A simple calculation leads to $x_k \sim 1/w_k$ which becomes an exact expression at $2 \leq k \ll B/2$, and gives the centres of the first stairs. For these values of x , one can show that \mathcal{D}_B is given by : $1 - 2k/B$, $k \geq 2$, in perfect agreement with numerical calculations. A simple parametrization of $\mathcal{D}_B(x)$ is then given by

$$x_k = 1 / \sum_{i=0}^{k-1} \binom{B}{i}, \quad k = 2, 3, \dots \quad (\text{centres of the stairs}) \tag{3}$$

$$\mathcal{D}_B(x_k) = 1 - 2k/B \quad (\text{height of the stairs}).$$

Note that, close to x_k , adding or removing a few points in Ω does not modify the hierarchy in an appreciable way, leading to the appearance of robust stairs. The result of equation (3) becomes more and more accurate at large B and for $k \ll B/2$. For $x \ll 1$, however, some deviations occur for various reasons. Beside trivial finite size effects, the most important source of deviations is the broadening of the probability distribution of nearest-neighbour distances. Indeed, at $x \ll 1$, different values of these distances (Fig. 3) become equiprobable, and lead to a fluctuation of distances of order $B^{1/2}$. Such fluctuations result in an overestimation of $\mathcal{D}_B(x)$ at small x , as given by equation (3).

5. Conclusion.

Our main conclusions are summarized in the introduction of this paper. The degree of ultrametricity \mathcal{D} , as introduced in this paper, may provide some insight in the geometrical structure of phase spaces. This is clearly shown in the worked example of section 4. It would be useful to extend such an analysis of the subdominant ultrametric, and the measure of the degree of ultrametricity \mathcal{D} to specific statistical-mechanical models or combinatorial optimization problems.

References

- [1] MÉZARD, M., PARISI, G., SOURLAS, N., TOULOUSE, G., VIRASORO, M. A., *J. Physique* **45** (1984) 843.
TOULOUSE, G., *Helv. Phys. Acta* **57** (1984) 459.
 - [2] See for instance BOURBAKI, N. : « Eléments de Mathématique », chap. 9, *Topologie générale*, p. 91.
Nouvelle édition 1974 (Diffusion CCLS, Paris).
 - [3] PARGA, N., PARISI, G., VIRASORO, M. A., *J. Physique Lett.* **45** (1984) 1063.
SOURLAS, N., *J. Physique Lett.* **45** (1984) 969.
LALLEMAND, P., DIEP, H. T., GHAZALI, A., TOULOUSE, G. (1985) Preprint.
 - [4] BACHAS, C. P., *Phys. Rev. Lett.* **54** (1985) 53.
KIRKPATRICK, S., TOULOUSE, G., *J. Physique* **46** (1985) 1227.
 - [5] Recall that ultrametrics refer to metric spaces where the known triangle inequality : $d(x, y) \leq d(x, z) + d(z, y)$ is replaced by the strongest condition $d(x, y) \leq \text{Max}(d(x, z), d(z, y))$. Note that these two inequalities are actually particular cases of the generalized triangle inequality : $d(x, y) \leq [d(x, z)^r + d(z, y)^r]^{1/r}$, where $r > 0$ denotes a real number. Metric spaces correspond to $r = 1$ and $r = \infty$ reproduces the ultrametric inequality.
 - [6] BENZECRI, J. P., *L'analyse des données : la taxonomie*, tome **1** (Dunod, Paris) 1980, 3^e édition.
JARDINE, N. and SIBSON, R., *Mathematical Taxonomy* (J. Wiley and Sons Ltd, London, New York) 1971.
 - [7] JARDINE, C. J., JARDINE, N. and SIBSON, R., *Math. Biosci.* **1** (1967) 173.
JARDINE, N. and SIBSON, R., *Math. Biosci.* **2** (1968) 465 ; *Comput. J.* **11** (1968) 177.
 - [8] For an excellent review, see MARTAGH, F., *Comput. J.* **26** (1983) 354.
 - [9] PAPADIMITRIOU, C. H. and STEIGLITZ, K., *Combinatorial Optimization* (Prentice Hall) 1982.
 - [10] See for instance SOLOMON, H., *Geometric probability* (Society for Industrial and Applied Mathematics, Philadelphia) 1978.
-