HIERARCHICAL DIFFUSION*

CONSTANTIN P. BACHAS**

Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305

ABSTRACT

We review the solution and properties of the diffusion equation in a hierarchical or ultrametric space [1].

One of the oldest equations in physics is the equation of diffusion. In its most general form it reads:

$$\frac{dP_i}{dt} = \sum_{j=1}^{N} \epsilon_{ij} P_j \tag{1}$$

where P_i is the probability of finding a particle at site i ($i = 1, ..., N$) of some arbitrary discretized space $\mathcal{M}$, and ϵ_{ij} is the hopping or transition probability per unit time, from site j to site i. This must of course be positive:

$$\epsilon_{ij} \geq 0 \qquad for \quad i \neq j \tag{2a}$$

and, in order to conserve total probability ($\sum_i P_i(t) = 1$), we must have :

$$\sum_{i=1}^{N} \epsilon_{ij} = 0 \tag{2b}$$

which fixes the diagonal elements ϵ_{ii} . Furthermore, I will assume in this talk that the transition matrix is symmetric:

$$\epsilon_{ij} = \epsilon_{ji} \tag{3}$$

i.e. that there is equal probability of hopping forward and backward between two sites. This restriction is not necessary for having a well defined diffusion problem, but as we will see it can be effectively lifted.

Invited talk at the meeting on
"Measures of Complexity", Rome, Italy, September 30 – October 2, 1987

Some general facts follow immediately from conditions (2-3). Firstly, there is a time-independent, steady-state solution, corresponding to equal probability at every site:

$$P_i(t) = \frac{1}{N} \qquad \text{for all } i \qquad (4)$$

Secondly, for any vector $\mathbf{x}$:*

$$\mathbf{x}\epsilon\mathbf{x} = \sum_{i \geq j} \epsilon_{ij}(2\mathbf{x}_i\mathbf{x}_j - \mathbf{x}_i^2 - \mathbf{x}_j^2) \leq 0 \qquad (5)$$

which shows that with the exception of the zero eigenvalue, corresponding to the vector (4), all other eigenvalues of the matrix ϵ are negative. Denote by

$$\epsilon^{(1)} = 0 > \epsilon^{(2)} \geq \; .... \; \geq \epsilon^{(N)} \qquad (6)$$

these eigenvalues, and by $\mathbf{v}_i^{(1)} = \frac{1}{\sqrt{N}}$, $\mathbf{v}^{(2)}$,...,$\mathbf{v}^{(N)}$ the corresponding orthonormalized eigenvectors. Knowledge of these clearly suffices to solve the diffusion problem (1), for any initial probability distribution $P(0)$ at $t = 0$. The result is:

$$P_i(t) = \frac{1}{N} + \sum_{I=2}^{N} < P(0)|\mathbf{v}^{(I)} > \mathbf{v}_i^{(I)} \; exp(-t/\tau^{(I)}) \qquad (7)$$

where we have here defined the characteristic times :

$$\tau^{(I)} = -\frac{1}{\epsilon^{(I)}} \quad , \qquad (8)$$

$< \mathbf{x}|\mathbf{y} >$ is the usual vector inner product, and we have used the fact that $< P(0)|\mathbf{v}^{(1)} > = \frac{1}{\sqrt{N}}$ since total probability is always one.

It follows easily from (7) that, at large t, one always approaches asymptotically the steady state (4). For finite spaces $\mathcal{M}$ the approach is exponential and dominated, generically, by the largest characteristic time $\tau^{(2)}$. For infinite spaces the relaxation can be slower than exponential, if there is a sufficient concentration of characteristic times at infinity. For instance if the density of characteristic times behaves asymptotically when $\tau \to \infty$ as:

$$\rho(\tau) \sim \tau^{-\nu-1} \qquad (9)$$

then for generic initial conditions the relaxation is described at large times by a power law:
$\int^{\infty} \rho(\tau)e^{-\frac{t}{\tau}}d\tau \sim t^{-\nu}$.

---

* We will often drop the vector index labelling the sites of $\mathcal{M}$; summation over repeated indices is as usual implied.

Now we all learned in highschool how to solve the problem of diffusion on a *regular lattice* in D-dimensional Euclidean space. The eigenvectors of the transition matrix are in this case plane waves, their eigenvalue is proportional to momentum squared, and $\nu = \frac{1}{2}D$. Another situation in which one can, if not completely solve the problem, at least calculate exponents such as ν, is when the space $\mathcal{M}$ is scale-invariant or *fractal* [2]. Here I will consider yet another kind of geometry, *ultrametric* geometry, which has neither translational nor , necessarily, scale invariance [3] . Ultrametricity is the statement that , for all i,j and k :

$$\epsilon_{ij} \geq min(\epsilon_{ik}, \epsilon_{jk}) \tag{10}$$

which implies that given any three sites, the two smallest transition rates are equal. A more descriptive but completely equivalent way of saying this is that the sites can be organized as the leaves of some genealogical tree, so that transition rates are only a (decreasing) function of relation, i.e. a) $\epsilon_{ij} = \epsilon_{A(i,j)}$ is only a function of the nearest common ancestor $A(i,j)$ of i and j , and b) ϵ_A is monotone decreasing as A moves up, along any path, towards the patriarch or root of the tree. We may in fact encode all the information about the transition matrix ϵ in the tree, by stretching the heights h_A of its branch-points , so that $\epsilon_A \equiv e^{-h_A}$ (see fig. 1a). The height $h_{A(i,j)}$ can be thought of as an energy barrier, through which a particle going from site i to site j must penetrate. With this convention, both topology and branch heights will be relevant when we refer to trees in the sequel.

Much of the interest in ultrametric geometry was spurred by the discovery [4] that it describes the space of spin-glass states in mean field theory [5] . An exact or approximate hierarchical organization, however, also characterizes a wide variety of other natural and artificial systems. For the mathematicians and, more recently the string theorists, a familiar example of an ultrametric norm is the p-adic norm on rational numbers. What is surprising is that ultrametricity is powerful enough to allow us to solve the problem of diffusion exactly, without any further assumption or approximation [1] .

In order to describe the diagonalization of an arbitrary ultrametric transition matrix, let us introduce some notation (see also fig.1a): for any branch point or tree leaf B, we let B_n be its unique n-th ancestor, N_B the total number of its final descendants , i.e. tree leaves generated by B, and $\chi^{(B)}$ their characteristic function:

$$\chi_i^{(B)} = \begin{cases} 1, & \text{if i is a descendant of B} \\ 0, & \text{otherwise.} \end{cases}$$

Then for each B :

$$\mathbf{u}^{(B)} = \frac{1}{N_B}\chi^{(B)} - \frac{1}{N_{B_1}}\chi^{(B_1)} \tag{11}$$

is an eigenvector of the transition matrix, that describes the process of equilibration
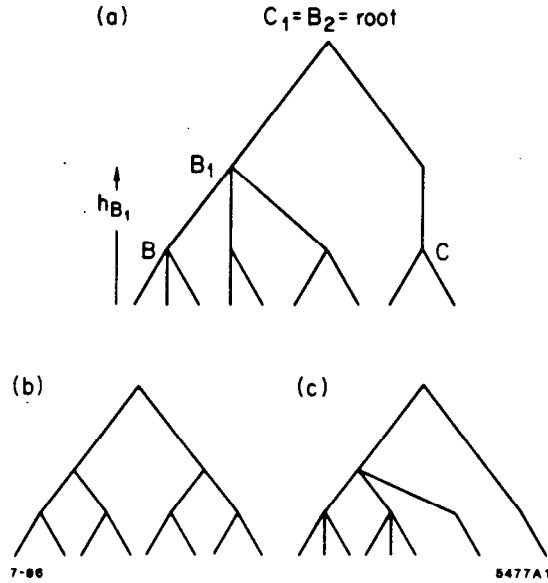
Figure 1 a) A generic tree illustrating our notation; the root is the father of B_1 and C, and the grandfather of B. The nodes B_1, B and C have 7, 3 and 2 final descendants respectively. All heights are measured from the leaves, which represent the sites of $\mathcal{M}$. b) A self-similar, uniformly bifurcating tree . c) A most complex tree leading to slowest possible relaxation; its silhouette is the same as for (b), since total population doubles at every generation.

between the descendants of B and those of all his brothers. The corresponding eigenvalue, or inverse characteristic time, of this process can be expressed as a sum over all ancestors of B:

$$\epsilon^{(B)} = \sum_{n=1}^{root} N_{B_n}(\epsilon_{B_n} - \epsilon_{B_{n+1}}) \tag{12}$$

where by abuse of notation *root* stands here for the number of generations between B and the tree root, and all terms in eqs. (11) and (12) that refer to ancestors of the root should , by convention, be dropped. Suppose now that at $t = 0$, the particle is with probability one at a given site, i.e. tree leaf, L. We may decompose this initial condition in terms of the eigenvectors (11) :

$$\delta_{i,L} = \sum_{n=0}^{root} \mathbf{u}^{(L_n)} \tag{13}$$

Since any other initial condition can obviously be written as a linear combination of (13), we have thus solved completely the problem of ultra-diffusion, *for any underlying tree.*

Let me now concentrate on the exponent ν that measures the speed of relaxation for infinite trees. Since there are many many more trees than real numbers, we expect some kind of

universality. In other words ν should only depend on very few characteristics of the tree. One obviously relevant characteristic is the asymptotic rate of population growth, or *silhouette* :

$$s = \lim_{h,\Delta h \to \infty} -\frac{\Delta log n(h)}{\Delta h} \qquad (14) .$$

where $n(h)$ is the population at height h. A large s means that there are on the average more sites available for hopping in, at given barrier height, and this should speed up relaxation. Thus it makes sense to fix the silhouette s, and ask how ν depends on the remaining characteristics of the tree. For instance one may want to compare relaxation on the trees of fig. 1b and 1c, which have the same silhouette, i.e. would look the same if they were to wear a coat, but have completely different internal structure. The following three theorems address this issue. They hold under the assumption of stable relaxation, meaning that the autocorrelation function never decays faster than exponentially in time. I state these theorems without proof, since detailed proofs can be found in ref. [1]:

Theorem 1: *For uniformly multifurcating trees, $\nu_{uniform} = \frac{s}{1-s}$. For any other tree, $\nu \leq \frac{s}{1-s}$.*

Examples of uniformly multifurcating trees are the tree of fig.1b, or the tree of p-adic numbers. The above result shows that they lead to optimal relaxation. The next result identifies a class of trees leading to the slowest possible relaxation:

Theorem 2: *For the tree of fig. 1c, $\nu = s$. For any other tree, $\nu \geq s$.*

The third and last result shows that structural noise is irrelevant, in that it modifies the power law decay of autocorrelations by, at most, logarithmic corrections:

Theorem 3: *For trees whose branching ratio at every node is an identically distributed, independent random variable, $\nu_{random} = \frac{s}{1-s}$.*

Now both uniform and uniformly random trees are self-similar structures, whose parts are on the average identical to the whole. Thus we may interpret the above results as saying that , for fixed s, the dynamic exponent $-\nu$ measures the lack of self-similarity, or the *complexity* of the hierarchical structure. One can in fact define other, static measures of a tree's complexity, that share the same qualitative features as $-\nu$ [6,7] . An example is the critical threshold for percolation, p_c, from the root to the bottom of the genealogical tree. It can be shown [7] that

p_c is also minimized by self-similar trees, is insensitive to noise, and is maximized by the very unbalanced tree of fig. 1c. This suggests that ν may in fact be a monotone function of p_c, but I have not been able to prove this.

Let me make here a parenthetical comment on semantics. The use of the word *complexity* in the above context can be motivated in many ways. For instance it is in accordance with our intuitive notion that complex is neither ordered nor random [6]. And the threshold for percolation on the tree of winning strategies of a game is, indeed, a measure of the fault-tolerance and hence of the complexity of the game [7] . Nevertheless, I am fully aware of the fact that *complexity* figures high in the list of most used and abused words in the scientific literature (to give a very banal example: complex analysis is as we all know much simpler than real analysis). For this reason some other term for ν and p_c might have been more appropriate.

I will conclude with some remarks about where and how ultradiffusion could be usefull [8]. Clearly, since an ultrametric transition matrix allows for infinite range hoppings, we do not expect it to describe diffusive processes in a finite- dimensional space. $\mathcal{M}$ could however be either an infinitely connected artificial system , like a neural network, or the configuration space of a spin- or other statistical mechanical model . Consider in particular the mean-field spin glass [5]. It is reasonable to assume that like thermodynamic equilibrium states, long-lived metastable states also have a hierarchical organization [4]. Since the appearance of ultrametricity is, however, in this case spontaneous, we do not know a priori the structure of the underlying tree, which we need as an input in the diffusion equation. We could however try to work backwards; indeed, the hopping rates between metastable states are given by: $\epsilon_{ij} = e^{-\Delta F_{ij}/T}$, where ΔF_{ij} are free energy barriers. Let us make the naive assumption that the distribution of these barriers does not change, or changes very slowly with temperature. Then clearly the silhouette s is proportional to temperature, and the temperature-dependent dynamic exponent behaves like:

$$\nu(T) = \begin{cases} \frac{T}{T_c - T} & \text{for self-similar trees} \\ T & \text{for most complex trees} \end{cases} \tag{15}$$

below the critical temperature T_c, at which relaxation becomes unstable [1]. Note in particular that the transition to instability ($\nu = \infty$) is continuous in the case of self-similar trees, and discontinuous for the most complex ones. Now in the mean-field spin glass it is known [9] that below T_c:

$$\nu = \frac{1}{2} - \frac{T - T_c}{\pi T_c} + o(T - T_c)^2 \tag{16}$$

so that the transition is discontinuous. Thus, if we take the above naive model seriously, we would be tempted to conclude that the tree of the mean field spin glass is not self-similar. Interestingly

enough, numerical efforts to reconstruct this tree give a result that looks very much like fig.1c [10].

Finally let me point out that the set of all ultrametric transition matrices sharing the same *topological tree*, is closed under both addition (trivially) and multiplication (we let the reader prove this for himself). Thus, both ultrametricity and the topology of the tree are stable under time-rescalings. It would be very interesting to study whether these stable trajectories in the space of all possible transition matrices, have any basins of attraction. In this case ultrametricity could be recovered as an effective property at large times. It would also be interesting to see whether the above ideas on complexity could be extended to the case of diffusion on multifractals [11].

REFERENCES

1. C.P.Bachas and B.A.Huberman , Phys.Rev.Lett. 57 (1986) 1965 ; J.Phys. A20 (1987) 4995 .

2. P.G. de Gennes, Recherche 7 (1976) 919 ; S.Alexander and R.Orbach, J.Physique Lett. 43 (1982) L625 ; R.Rammal and G.Toulouse, ibid 44 (1983) L13 .

3. Scale invariant ultradiffusion was introduced and studied before ref. [1], in different contexts and variations , by many authors : B.A.Huberman and M.Kerszberg, J.Phys. A18 (1985) L331 ; S.Teitel and E.Domany, Phys.Rev.Lett. 55 (1985) 2176 and 56 (1985) 1755 ; A.Maritan and A.L.Stella, ibid 56 (1986) 1754 and J.Phys. A19 (1986) L269; S.Grossman, F.Wegner and K.H.Hoffmann, J.Physique Lett. 46 (1985) L575 ; G.Paladin, M.Mezard and C.De Dominicis, ibid 46 (1985) L985 ; M.Schreckenberg, Z.Phys. B60 (1985) 483 ; A.T.Ogielski and D.L.Stein, Phys.Rev.Lett. 55 (1985) 1634 ; D.Kumar and S.R.Shenoy, Solid State Comm. 57 (1986) 927 ; A.Blumen, J.Klafter and G.Zumofen, J.Phys. A19 (1986) L77 .

4. M.Mezard, G.Parisi, N.Sourlas, G.Toulouse and M.Virasoro, Phys.Rev.Lett. 52 (1984) 1156 and J.Physique 45 (1984) 843 ; for a review of ultrametricity see also R.Rammal, G.Toulouse and M.A.Virasoro, Rev.Mod.Phys. 58 (1986) 765 and references therein .

5. D.Sherrington and S.Kirkpatrick, Phys.Rev.Lett. 35 (1975) 1792 .

6. B.A.Huberman and T.Hogg, Physica 22D (1986) 376 ; B.A.Huberman and H.A.Ceccatto, Xerox PARC preprint (1987) .

7. C.P.Bachas and W.F.Wolff, J.Phys.A20 (1987) L39 .

8. For a discussion of this issue see : R.G.Palmer, in Proceedings of the Heidelberg Colloquium on Glassy Dynamics and Optimization, eds. J.L.van Hemmen and I.Morgenstern, Springer Verlag 1986 and references therein, and also ref.[3].

9. H.Sompolinsky and A.Zippelius, Phys.Rev.Lett 47 (1981) 359 and Phys.Rev. B25 (1982) 6860 .

10. N.Parga, CERN-TH 4410 preprint (1986) .

11. G.Paladin and A.Vulpiani, Anomalous Scaling Laws in Multifractal Objects, Physics Reports, to appear.