Spectral signatures of hierarchical relaxation

B. Gaveau
Université Paris 6, Géométrie des équations aux dérivées partielles
et physique mathématique
Case 172, 4, Place Jussieu, 75252 Paris Cédex 05, France

A. Lesne
Université Paris 6, Laboratoire de Physique Théorique des Liquides
Case 121, 4, Place Jussieu, 75252 Paris Cédex 05, France

L. S. Schulman
Physics Department, Clarkson University, Potsdam, NY 13699-5820, U.S.A., and
Physics Department, Technion, Haifa, Israel

Physical Letters A **258**, 222–228 (1999)

Using dynamical concepts of phase transitions developed in earlier work, we exhibit the characteristic features of the spectrum and eigenfunctions to be expected when a hierarchy of phase transitions is present in a system, such as is expected to occur for spin glasses.

1. Introduction

The concept of phase transition is most precise in the context of equilibrium statistical mechanics in the thermodynamic limit. Unfortunately many systems of interest are excluded by the narrowness of this definition. In this paper we continue our master equation approach [1,2] to this problem and extend that method to structures having strong resemblance to those believed to exist in spin glasses. We remark that for spin glasses the perceived need to take a thermodynamic limit has vastly complicated the enterprise and that a dynamical approach, which is the essence of our method, is being pursued by other workers in this field as well [3].

The master equation approach is applicable to both traditional physical systems and to others of a more general nature [4,5]. In previous publications [1] we used this framework to provide concepts of entropy, dissipation, currents, and fluctuation-dissipation theorems, with or without detailed balance. In [2] we presented a concept of first order phase transition from this perspective. In the present paper we accommodate the hierarchical structure often attributed to spin glasses within the master equation approach, and show how spectral properties (especially *left* eigenfunctions) reflect the physically significant structure. We comment that our treatment of phase transitions has two principal differences from other approaches. First, it is dynamical, letting the system define its own metastability. Second, it is not infinitely sharp, in the sense that only asymptotic statements are made (no thermodynamic limit is taken). Arguments for these differences have already been made in the relatively simpler case of ordinary (macroscopic) metastability [6].

In Sec. 2 we outline our approach, with emphasis on phase transitions. Following that we comment on systems with slower than exponential relaxation and the implications for the transition matrix spectrum. In Sec. 3 a dynamical distance function is introduced and used to define a coarse graining. Finally in Sec. 4 the uses of our approach for spin-glasses are presented.

2. General framework

Dynamics are expressed through a stochastic matrix giving rates for transitions among the microscopic states. The latter are assumed to be elements of a finite set, X. Thus R is defined by $R_{xy} \equiv \Pr(x \leftarrow y)$ in one time step. We assume R to be irreducible. For many purposes these restrictions are not essential. Entropy, relative entropy, a fluctuation dissipation theorem and other results are discussed in [1].

The eigenvalues of R are numbered in order of decreasing magnitude, starting from the label 0. The corresponding left and right eigenvectors are denoted A_k and p_k respectively ($k = 0, 1, \ldots$). (Although R need not have a spectral expansion, in this article we will for convenience use such a representation. All we actually require, however, is the existence of eigenvectors for the eigenvalues near unity.) By probability conservation $\sum_x R_{xy} = 1$ (stochasticity). It follows that R has an eigenvalue $\lambda_0 = 1$ (which on general grounds is the largest possible magnitude) and left eigenvector that can be taken to be $A_0(x) \equiv 1$. The corresponding right eigenvector, p_0, has all elements strictly positive and is the stationary probability distribution.

In [2], for a transition involving two phases only, we begin with the assumption that λ_1 is strictly less than 1 and non-degenerate. Moreover, the existence of the transition is related to quasi-degeneracy through the assumption that there exists a range of times (t) such that $1 \sim \lambda_1^t \gg |\lambda_2|^t$. Intuitively, λ_2 and all smaller eigenvalues provide relaxation *within* a phase, while λ_1 governs the slower relaxation between the two phases.

From these assumptions we show that R is nearly reducible in the sense that the states on which p_0 has most of its weight break into two subsets—to be identified with the phases—on each of which R is almost stochastic. The key step in the proof is to show that the *left* eigenvector corresponding to λ_1 is "essentially" constant on each phase, so that it plays a role analogous to A_0 on the total space. To reach that conclusion we use the eigenvectors A_1 and p_1 associated with λ_1 and the orthogonality relations $\langle A_k | p_j \rangle = \delta(k, j)$. It follows that A_1 has a strictly positive maximum (A_M) and strictly negative minimum (A_m). These are attained on points y_M and y_m. We next define probability distributions p_M^t (p_m^t) as the distributions arising from a system's starting in y_M (y_m) at time 0. (Thus $p_M^t(x) = (R^t)_{x\, y_M}$.) Hence p_M^t and p_m^t can be interpreted as restricted stationary probabilities on the phases. The basis for most of our results is [2]

$$1 - \lambda_1^t = \sum_x p_M^t(x) \left(1 - \frac{A_1(x)}{A_M}\right) = \sum_x p_m^t(x) \left(1 - \frac{A_1(x)}{A_m}\right) \tag{1}$$

This says that so long as λ_1^t is close to 1, and for x such that either $p_M^t(x)$ or $p_m^t(x)$ is significant (i.e., x is in one of the phases), then the associated $A(x)$ cannot depart too much from its extreme value, A_M or A_m. This result can be generalized to the case of many phases by assuming that the first p eigenvalues of R are real and satisfy $\lambda_1^t, \ldots, \lambda_p^t = \mathrm{O}(1 - \varepsilon)$, for small ε. (Thus Eq. (1) does not depend on the smallness of λ_2.) To find the analogs of A_M and A_m and the associated points in X one uses convexity arguments in $\mathbf{R}^p$. We again identify $p + 1$ special X-space points y_q^*, $q = 1, \ldots, p + 1$, such that the following equalities hold

$$1 - \lambda_k^t = \sum_x p_{y_q^*}^t(x) \left(1 - \frac{A_k(x)}{A_k(y_q^*)}\right) \tag{2}$$

This holds *both for all* $k = 1, \ldots p$ *and for all* $q = 1, \ldots p+1$ [7]. In this way the basic phase transition breakup, effective reduction of R, etc., is carried over to the multiphase case.

Critical slowing down and other slowing down

Recalling the spectral expansion, $R^t = \sum_\alpha \lambda_\alpha^t p_\alpha A_\alpha$ [8], and the identity

$$\Gamma(\eta)\, t^{-\eta} = \int_0^\infty du\, u^{\eta-1} e^{-ut}$$

(with $\lambda \sim e^{-u}$), it is clear that to get power law dropoff in some property it is necessary to have a bunching of eigenvalues near 1. However, such bunching does not by itself characterize critical points. For example, diffusion in a volume of side L shows both slow relaxation and clustered spectrum for $L \to \infty$, but one would not ordinarily call this critical. On the other hand, the percolitis [9] phase transition has many properties that are critical, but its spectrum, which scales like $1/\sqrt{N}$ (with N the system size), strongly resembles that of diffusion in 4 dimensions. Our expectation is that non-mean field critical points have, in addition to eigenvalue clustering, an interplay of two distance concepts, one emerging from the structure of the points, $x \in X$, and the other induced by R (about which more later).

There are other kinds of slowing down. Examples are stretched exponentials, spin-glass relaxation and the decay of superfluid current [10]. In this article we consider the spin-glass case.

3. Coarse graining and dynamic distances

Both conceptually and practically the microscopic detail implicit in a state "x" is often excessive. We therefore define a coarse graining of X. This is done by means of a dynamic distance function used to sort points for combining in a single grain (or fiber).

The state X is divided into grains, labeled $\tilde{x} \in \widetilde{X}$. We write $x = (\tilde{x}, u)$ to indicate that a point of the original X is specified by its coarse grain label "$\tilde{x}$" and by a point u within that grain. In our coarse graining the stationary distribution p_0 plays a central role. The intention is that the slowest modes should survive the coarse graining. For $x = (\tilde{x}, u)$ and $y = (\tilde{y}, v)$, let $R(\tilde{x}, u; \tilde{y}, v) \equiv R_{xy}$ and define

$$\tilde{R}_{\tilde{x}\tilde{y}} = \sum_{u,v} R(\tilde{x}, u; \tilde{y}, v) p_0^c(\tilde{y}, v) \quad \text{with} \qquad p_0^c(\tilde{y}, v) \equiv \frac{p_0(\tilde{y}, v)}{\sum_{v'} p_0(\tilde{y}, v')}\,.$$

The new stationary distribution is $\tilde{p}_0(\tilde{x}) = \sum_u p_0(\tilde{x}, u)$. The currents $\widetilde{J}_{\tilde{x}\tilde{y}}$ [11] are likewise sums of the separate point-to-point currents within each grain. In principle this process need only leave the eigenvalue 1 intact, although for a sensible choice of grains the slowest eigenvalues are little affected. This is illustrated below.

In our abstract dynamical formulation there does not necessarily exist a concept of intrinsic distance on X related to the underlying space on which the microstates (x) are defined. All that is available is R, and that is where the distance comes from. We have considered several possibilities. From the two phase example one expects the first nontrivial left eigenvector to be a candidate for

distance, i.e., "Left eigenvector distance" $\equiv$ L.E.Dis.$(x,y) = |A_1(x) - A_1(y)|$. However, with many phases this is no longer so obvious a candidate. It also has a dependence on the total weight in the phase that may not be useful.

The definition we use preserves certain desirable ambiguities, in particular the time scale over which one wishes to coarse grain. We define

$$\mathrm{dis}_T(x,y) = \sum_u |\, p_x^T(u) - p_y^T(u)| \;, \quad \text{with} \;\; p_x^T(u) = (R^T)_{ux} \tag{3}$$

This means that you start the system alternatively in x or y and check at a later time T the extent to which the distributions overlap. For the two-phase case this will be closely related to our earlier tentative definition, but (3) is more generally useful. Other possibilities involve generalizations of "L.E.Dis.", using all A_k with the corresponding λ_k greater than some number $\tilde{\lambda}$ (close to 1, say). The selection of a $\tilde{\lambda}$ would correspond to the choice of T. We remark that distance concepts similar to those we introduce here have appeared in many contexts and the the use of "slow" variables of one form or another leads to a variety definitions (and some of these have as their purpose coarse graining, along the lines we take below). A sampling of references is [12, 13, 14].

The notions of coarse graining and distance are naturally combined by using the distance function to define grains. We give an example using a two phase system that we now construct. Let Q be a stochastic matrix, created by first generating a random matrix with each element selected uniformly on $[0,1]$. Then divide each column by the sum of its elements. We combine four such $N\times N$ matrices to produce a $2N\times 2N$ matrix:

$$\widehat{R} = \begin{pmatrix} Q(1,1) & \varepsilon_1 Q(1,2) \\ \varepsilon_2 Q(2,1) & Q(2,2) \end{pmatrix} \tag{4}$$

Each column of $\widehat{R}$ is then divided by the sum of the elements in that column to produce a stochastic matrix, R, that has the properties of a two-phase system, provided ε_1 and ε_2 are small. For $1 \gg \varepsilon_1 \gg \varepsilon_2$ (> 0) one of the phases is metastable; when the εs are equal the phases get equal weight. Generically this is *not* an equilibrium situation since nonzero currents are present in the stationary distribution.

We coarse grain this system in stages. The distances between each pair of states for a particular time T is calculated. The two points with smallest relative distance are combined in a single grain. Then the two nearest of the remaining states are combined, etc., halving the space (for convenience N is a power of 2 times a small prime). A reasonable value of T depends on the εs. This process is repeated until the final matrix is small. Generally the following structure is preserved:

$$R \sim \begin{pmatrix} \text{big} & \text{small} \\ \text{small} & \text{big} \end{pmatrix}$$

For a specific example, we show in Table 1 the values of the two largest nontrivial eigenvalues. The smaller of these represents the relaxation time *within* a phase. What makes this work is that for the specific T (of Eq. (3)), the points within each phase relax completely, but do not spill over to the other phase. Thus they are preferentially combined at each stage of coarse graining. The

slow process, the transition between phases, is not affected by this unification. Hence the largest nontrivial eigenvalue is unchanged.

In Fig. 1 we show both the stationary probability distributions (p_0) and first nontrivial left eigenvectors (A_1) at various stages of coarse graining. Because ε_1 is considerably different from ε_2 the stationary weights of the two phases are different. The approximate constancy of the left eigenvectors is also evident [15]. Note that where p_0 is large, A_1 takes one value, where it is small, the other. (The ability to see this is an advantage of taking $\varepsilon_1 \neq \varepsilon_2$.)

4. Hierarchical dynamical first order phase transitions

Relaxation in spin-glasses is notoriously slow. Using traditional criteria, the issue of whether or not a phase transition exists is complicated by the problems of defining the thermodynamic limit. A picture that has won acceptance is a hierarchical structure in which several states of relatively low free energy can, by a massive overturning of spins, go to a yet lower free energy, with this pattern repeating at many scales. We will not try to summarize the state of knowledge in this field, but only display a hierarchical structure within a master equation context that is suggestive of the situation just described. This structure gives a *dynamical* characterization of the spin-glass transition. Another, simpler situation where the known properties of the model suggest hierarchical structure is in multicomponent one-dimensional driven diffusive systems [16].

As an example, we consider a multiphase system such that the number of independent phases depends on the time scale for observation. On the shortest time scale the phases consist of n_1 states (for simplicity all phases have the same number of states). For larger t, n_2 such groups of states flow back and forth among themselves, but do not go (or seldom go) anywhere else. As time increases, ever larger groups are connected.

We study such a system with a particular matrix and parameters. What will be seen is that *the left eigenvectors are locally constant on the phases*, with the degree of "constancy" related to the time scale for that phase. Finally we will examine the question of whether this method could lead to a way to deal with the enormous transition matrices that would arise in even a small 3-dimensional spin glass.

The matrix is generated recursively. At the lowest level is a collection of $n_1 \times n_1$ stochastic matrices. These are coupled weakly, as in Eq. (4), except that instead of only two, any finite number, n_2, is allowed. These larger units are again grouped, n_3 at a time, with weaker coupling. The model is thus defined by the sequences $n_1, n_2, \ldots$, and $\varepsilon_1, \varepsilon_2, \ldots$ (the couplings). In our example the hierarchy ends at n_3. Our choice was $\{n\} = (10, 2, 2)$ and $\{\varepsilon\} = (0.01, 0.0001)$. (Everything works for larger systems, but more figures would be needed.) In Fig. 2 we show the first 4 non-trivial left eigenvectors. The absolute values of the associated eigenvalues are: $0.9999, 0.9921, 0.9907, 0.1971$, showing there to be three distinct time scales: $1/\log\lambda_1$, $1/\log\lambda_2$, ($\approx 1/\log\lambda_3$) and $1/\log\lambda_4$. As described in the figure caption there is a clear hierarchy of phases, with the shorter lived ones subsets of the longer lived (eigenvalue closer to 1) ones. The right eigenvectors do show similar behavior in the sense that substantial departure from zero occurs only on ranges of relative constancy of the left eigenvectors. However, the structure is more difficult to discern because the actual value assumed by the right eigenvectors is not systematically identifiable with a particular phase. Note too that the left eigenvectors can assume complex values, a situation not explicitly considered in [2].

We next take up coarse graining. A single coarse graining, by a factor 2, leaves the entire structure intact. We will not repeat the figure corresponding to Fig. 2, since it has identical features, except for the abscissa running from 1 to 20, instead of 1 to 40. The new set of eigenvalues is $0.9999, 0.9921, 0.9907, 0.2114$, virtually identical, except for a relative slowdown in internal equilibration within each phase. The next coarse graining, to a 10×10 matrix, does have an interesting complication due to 40 not being divisible by $2^2 n_2 n_3$. We show the left eigenvectors for this case in Fig. 3.

As remarked earlier, coarse graining has practical as well as conceptual value, especially for the numerical use of the master equation formulation. Thus the number of spins in 3 dimensional cubes of sides 2 and 3 are 8 and 27, respectively. The sizes of the state spaces for even so small a system are $2^8 = 256$ and $2^{27} \sim 10^8$, leading in the latter case to impossibly large matrices. The suggestion that comes to mind is a dynamical renormalization based on the distance concept introduced earlier. Thus the 256 states of a $2\times2\times2$ Ising model cube would be allowed to evolve under stochastic dynamics (the coupling could be ferromagnetic, random, etc.) and distances of the resulting distributions evaluated (cf. Eq. (3)). If the resulting states fall into a relatively small number of classes, and if the results were not sensitive to either the boundary conditions or the particular random, quenched coupling (in the spin glass case), these classes and dynamically calculated (from simulations) transition probabilities among them could themselves be used for the next stage of scaling.

In any case, whether or not explicit transition matrices can be used for realistic models, the difficulties of using traditional thermodynamic limit ideas suggests the value of another theoretical framework. For example, we have already used this framework in mean field models such as percolitis [9] and one may expect analogous methods to work for mean field spin glass models.

Acknowledgements This work was supported in part by the United States National Science Foundation grant PHY 97 21459.

References

1. B. Gaveau and L. S. Schulman, Master equation based formulation of non-equilibrium statistical mechanics, J. Math. Phys. **37**, 3897 (1996); B. Gaveau and L. S. Schulman, A general framework for non-equilibrium phenomena: The master equation and its formal consequences, Phys. Lett. A **229**, 347 (1997).

2. B. Gaveau and L. S. Schulman, Theory of non-equilibrium first order phase transitions for stochastic dynamics, J. Math. Phys. **39**, 1517 (1998).

3. C. M. Newman and D. L. Stein, Equilibrium pure states and nonequilibrium chaos, J. Stat. Phys., to appear (1999); C. M. Newman and D. L. Stein, Blocking and persistence in the zero-temperature dynamics of homogeneous and disordered Ising models, preprint.

4. As an example far removed from the physical sciences, see E. O. Wilson, *Naturalist*, Island Press, Washington (1994), p. 311, where a discussion of the work of Stuart Altman is given.

5. M. Cieplak, M. Henkel, J. Karbowski and J. R. Banavar, Master equation approach to protein folding and kinetic traps, Phys. Rev. Lett. **80**, 3654 (1998).

6. L. S. Schulman, System-size Effects in Metastability, pp. 490-518, in *Finite Size Scaling and Numerical Simulation of Statistical Systems*, V. Privman, ed., World Scientific, Singapore (1990).

7. Eq. (2) is a single component of Eq. (9.7) in [2].

8. Even when there is no spectral expansion, small changes in the dynamics can generally recover this property. Similarly, eigenvalues (other than the zeroth) can be forced to have magnitude less than 1 by adding a small multiple of the identity matrix to the dynamical matrix and shrinking appropriately to maintain stochasticity.

9. L. S. Schulman and P. E. Seiden, Percolation and Galaxies, Science, **233**, 425 (1986); B. Gaveau and L. S. Schulman, Finite-size scaling for mean-field percolation, J. Stat. Phys. **70**, 613 (1993); B. Gaveau and L. S. Schulman, Fluctuations in mean field self organized criticality, J. Stat. Phys. **74**, 607 (1994); V. Privman and L. S. Schulman, Infinite-Range Mean-Field Percolation: Transfer Matrix Study of Longitudinal Correlation length, J. Stat. Phys. **64**, 207 (1991).

10. T. G. Wang and I. Rudnick, Anomalous Attenuation of Third Sound, J. Low Temp. Phys. **9**, 425 (1972).

11. The current is defined as $J_{xy} \equiv R_{xy}p_0(y) - R_{yx}p_0(x)$. This is nonzero only in the absence of detailed balance, and is thus characteristic of nonequilibrium situations.

12. M. C. Mackey, *Time's Arrow: The Origins of Thermodynamic Behavior*, Springer, New York (1992).

13. G. Nicolis, S. Martinez and E. Tirapegui, Chaos, Solitons and Fractals **1**, 25-37 (1991).

14. P. Diaconis and M. Shashahani, SIAM J. Math. Anal. **18**, 208-218 (1987); P. Diaconis and D. Stroock, Ann. Applied Prob. **1**, 36-61 (1991).

15. For states of small weight A_1 can take intermediate values. This occurs in realistic models, where most microscopic states are neither one phase nor the other. In a one-dimensional model, say diffusion in a potential, the phases would be states in the minima, and the intermediate values of A_1 would occur in the barrier region.

16. M. R. Evans, Y. Kafri, H. M. Koduvely, D. Mukamel, (1998): Phase separation and coarsening in one-dimensional driven diffusive systems. Phys. Rev. Lett. **80**, 425-429 (1998); M. R. Evans, Y. Kafri, H. M. Koduvely, D. Mukamel, Phase separation and coarsening in one-dimensional driven diffusive systems: Local dynamics leading to long-range Hamiltonians. Phys. Rev. E **58**, 2764-2778 (1998).

Table

Table 1. Two largest eigenvalues (excluding the trivial value, 1) for a succession of coarse grainings. If an eigenvalue is complex, its magnitude is given. The original matrix is 96 by 96 and the couplings are 0.01 and .001. The time used for distance measurements is 8. When the coarse graining has reduced the system to a 3-by-3 matrix, the two phase-two state description is no longer valid. 8 Matrix size 96 48 24 12 6 3 Largest 0.9891 0.9891 0.9891 0.9891 0.9891 0.6542

Second largest 0.0895 0.0657 0.0441 0.0274 0.0159 0.0126

Figure Captions

Figure 1. The upper row is the vector A_1 as a function of state label (which is arbitrary). The lower row is the stationary distribution, p_0 with the same labels. The first column shows the original 96-by-96 matrix. The second column shows the eigenvectors after two coarse grainings (each by a factor 2) and the third column after another 2. The last column brings the matrix down to 3-by-3, at which point the two-phase aspect is lost. Parameter values for the matrix are the same as in Table 1.

Figure 2. Absolute values of the first four nontrivial left eigenvectors for the matrix described in the text. The first such eigenvector ("no. 2") shows the system to have essentially two phases, while the next two show how those phases have internal two-phase structure. The next eigenvector, whose eigenvalue is distant from 1, is far from constant. Note though that its interesting behavior is confined to a single one of the previous phases.

Figure 3. Absolute values of the first four nontrivial left eigenvectors for coarse graining by a factor 4. The transition matrix is the same as that considered in Fig. 2.