Metastability and small eigenvalues in Markov chains

2000 J. Phys. A: Math. Gen. 33 L447

(http://iopscience.iop.org/0305-4470/33/46/102)

J. Phys. A: Math. Gen. **33** (2000) L447–L451. Printed in the UK PII: S0305-4470(00)15952-9

LETTER TO THE EDITOR

Metastability and small eigenvalues in Markov chains

Anton Bovier†, Michael Eckhoff‡, Véronique Gayrard§ and Markus Klein‡

† Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, D-10117 Berlin, Germany

‡ Institut für Mathematik, Universität Potsdam, Am Neuen Palais 10, D-14469 Potsdam, Germany

§ DMA, EPFL, CH-1021 Lausanne, Switzerland, and Centre de Physique Théorique, CNRS, Luminy, Case 907, F-13288 Marseille, Cedex 9, France

E-mail: bovier@wias-berlin.de, eckhoff@math.uni-potsdam.de, Veronique.Gayrard@epfl.ch and mklein@math.uni-potsdam.de

Received 26 July 2000

Abstract. In this Letter we announce rigorous results that elucidate the relation between *metastable states* and *low-lying eigenvalues* in Markov chains in a much more general setting and with considerably greater precision than has so far been available. This includes a *sharp* uncertainty principle relating all low-lying eigenvalues to mean times of metastable transitions, a relation between the support of eigenfunctions and the attractor of a metastable state and sharp estimates of the convergence of the probability distribution of the metastable transition times to the exponential distribution.

1. Introduction

The phenomenon of *metastability* has been a fascinating topic of non-equilibrium statistical mechanics for a long time. Recently, it has found renewed interest in the investigation of *glassy systems* and *ageing phenomena*, which appear to play a central role in many physical and non-physical systems. An approach to link metastability to spectral characteristics, in particular *low-lying eigenvalues* and the corresponding eigenfunctions, has been proposed by Gaveau and Schulman [9]. Such an approach is appealing not only because it allows us to characterize metastability in terms that are intrinsically dynamic and make no *a priori* reference to geometric concepts such as 'free energy landscapes', but also since it allows numerical investigations of metastable states via numerical spectral analysis (see Schütte *et al* [13, 14] for applications to conformational dynamics of biomolecules).

Relating metastability to spectral characteristics of the Markov generator or transition matrix is in fact a rather old topic. The earliest mathematical results go back at least as far as Wentzell [16] (see also [10] for more recent results) and Freidlin and Wentzell (see [7]). Freidlin and Wentzell relate the eigenvalues of the transition matrix for Markov processes with exponentially small transition probabilities to exit times from 'cycles'. Scoppola [15] gives a similar result based on a different renormalization procedure developed in [11, 12]. Wentzell has a result for the spectral gap in the case of certain diffusion processes. All these relations result in some way from the application of large-deviation methods and are, consequently, on the level of logarithmic equivalence, i.e. of the form $\lim_{\epsilon\downarrow 0} \epsilon \ln(\lambda_i^\epsilon T_i^\epsilon) = 0$ where ϵ is the small parameter, and λ_i^ϵ, T_i^ϵ are the eigenvalues and exit times, respectively. This rather crude level of precision persists also in the more recent literature and prevents, in particular, applications to systems with unbounded numbers of metastable states, which are particularly relevant for glassy systems.

In this Letter we announce results that—for a large class of Markov chains—improve this situation considerably: in particular we allow for the number of metastable states to grow (with, for example, the 'volume'), and we give precise control of error terms for 'finite-volume' systems. Moreover, we provide representations for all quantities concerned that are computable in terms of certain 'escape probabilities' that are in turn well controllable through variational representations [2].

Our starting point will be the definition of a *metastable set*, $\mathcal{M}_N$, of points, each of which is supposed to be a representative of one *metastable state*, on a chosen timescale. It is important that our approach allows us to consider the case where the cardinality of $\mathcal{M}_N$ depends on N. The key idea behind our definition will be that it ensures that the time it takes to visit the representative point once the process enters a 'metastable state' is very short compared to the lifetime of the metastable state. Thus, observing the visits of the process at the metastable set suffices largely to trace the history of the process.

The results presented here cover the case of what we would call generic metastability, i.e. each point of the metastable set corresponds to a different timescale and an associated isolated eigenvalue of unit multiplicity. Our methods are, however, also intended to treat more general situations where groups of finitely, or possibly countably many, states communicate on the same timescale. However, the analysis of the resulting phenomena is by no means trivial and requires a case by case analysis. As first example of a situation with an unbounded number of effectively communicating metastable states we refer to the analysis of ageing phenomena in the random energy model in [1].

A more detailed exposition of our results, as well as the proofs, will be given in two forthcoming papers [3, 6].

2. Metastable sets and metastable states

We will consider in the following Markov chains X_t with state space Γ_N, discrete time† $t \in \mathbb{N}$ and transition matrices P_N. We will assume that for any fixed N they are ergodic, and have a unique invariant distribution $\mathbb{Q}_N$. The special case when P_N is self-adjoint with respect to the measure $\mathbb{Q}_N$ is referred to as *reversible* dynamics. Some of our results will be sharper in that case. We are interested in the situation when these chains exhibit 'metastable' behaviour; loosely speaking, this means that the state space Γ_N can be decomposed into subsets $S_{N,i}$ such that the typical times the process takes to go from one such set to another are much larger than the time it takes to 'look like' being in equilibrium with respect to the conditional distribution $\mathbb{Q}_N(\cdot|S_{N,i})$. Some reflection shows that this statement has considerable difficulties and cannot be interpreted literally, and that a precise definition of metastability is a rather tricky business (see, for example, the recent discussion in [4]). We will give a precise definition that is, however, inspired by this vague consideration. The main point here is that one should make precise the two timescales we alluded to. We will take the following attitude: to appear ergodic within $S_{N,i}$, the process should have at least enough time to reach the 'most attractive' state within $S_{N,i}$, while at least the time to go from two such states in different metastable regions should be long compared to that time. Note that this concept is rather flexible and allows us, in general, to define metastable states corresponding to different timescales.

The following definition of 'metastable sets' follows this ideology; however, we prefer to use certain probabilities rather than actual times as criteria, mainly because these are more

† All results apply, however, also to continuous time.

readily computable. Linking them in a precise manner to times will be part of our results. We will write τ_I^x, for $x \in \Gamma_N$, $I \subset \Gamma_N$, for the first non-zero time at which the process started at x arrives at I.

Definition 2.1. *A set* $\mathcal{M}_N \subset \Gamma_N$ *will be called a set of* metastable points *if it satisfies the following conditions. There are finite positive constants* a_N, b_N, c_N *and* r_N *satisfying for some sequence* $\varepsilon_N \downarrow 0$, $a_N^{-1} \leqslant \varepsilon_N b_N$, *such that the following hold.*

(i) For all $z \in \Gamma_N$,

$$\mathbb{P}[\tau^z_{\mathcal{M}_N} < \tau^z_z] \geqslant b_N. \tag{2.1}$$

(ii) For any $x \neq y \in \mathcal{M}_N$,

$$\mathbb{P}[\tau^x_y < \tau^x_x] \leqslant a_n^{-1}. \tag{2.2}$$

(iii) We associate with each $x \in \mathcal{M}_N$ *its local valley*

$$A(x) \equiv \{z \in \Gamma_N : \mathbb{P}[\tau^z_x = \tau^z_{\mathcal{M}_N}] = \sup_{y \in \mathcal{M}_N} \mathbb{P}[\tau^z_y = \tau^z_{\mathcal{M}_N}]\}. \tag{2.3}$$

Then

$$r_N \geqslant \frac{\mathbb{Q}_N(x)}{\mathbb{Q}_N(A(x))} \equiv R_x \geqslant c_N^{-1}. \tag{2.4}$$

We will also write $T_{x,I} \equiv \mathbb{P}[\tau^x_I \leqslant \tau^x_x]^{-1}$. An important characteristic of the sets $I \subset \mathcal{M}_N$ is $T_I \equiv \sup_{x \in \mathcal{M}_N} T_{x,I}$. A simplifying assumption, that will be seen to ensure sufficient 'spacing' of the low-lying eigenvalues is that of 'genericity', defined as follows.

Definition 2.2. *We say that our Markov chain is* generic *on the level of the set* $\mathcal{M}_N$ *if there exists a sequence* $\epsilon_N \downarrow 0$, *such that the following hold.*

(i) For all pairs $x, y \in \mathcal{M}_N$, *and any set* $I \subset \mathcal{M}_N \backslash \{x, y\}$ *either* $T_{x,I} \leqslant \epsilon_N T_{y,I}$ *or* $T_{y,I} \leqslant \epsilon_N T_{x,I}$.

(ii) There exists $m_1 \in \mathcal{M}_N$ *such that for all* $x \in \mathcal{M}_N \backslash m_1$, $\mathbb{Q}_N(x) \leqslant \epsilon_N \mathbb{Q}_N(m_1)$.

Each of the elements of $\mathcal{M}_N$ in the generic case will then correspond indeed to a metastable state. Our first task is to identify precisely the notion of the exit time from a metastable state. To do so, we define for any $x \in \mathcal{M}_N$ the set $\mathcal{M}_N(x) \equiv \{y \in \mathcal{M}_N : \mathbb{Q}_N(y) > \mathbb{Q}_N(x)\}$; these are the points that are even more stable than x. The metastable exit time, $t_x \equiv \tau^x_{\mathcal{M}_N(x)}$, from x is then defined as the time of the first arrival from x in $\mathcal{M}_N(x)$. With this notion we can formulate our main result.

Theorem 2.3. *Assume that* $\mathcal{M}_N$ *is a metastable set and that the genericity assumptions are satisfied with* ε_N *such that* $r_N c_N \epsilon_N \downarrow 0$ *and* $\epsilon_N |\Gamma_N| |\mathcal{M}_N| \downarrow 0$. *Then the following hold.*

(i) For any $x \in \mathcal{M}_N$,

$$\mathbb{E}\tau_x = R_x^{-1} T_{x,\mathcal{M}_N(x)}(1 + o(1)). \tag{2.5}$$

(ii) For any $x \in \mathcal{M}_N$, *there exists an eigenvalue* λ_x *of* $1 - P_N$ *that satisfies*

$$\lambda_x = \frac{1}{\mathbb{E}\, t_x}(1 + o(1)). \tag{2.6}$$

Moreover, in the reversible case there exists a constant $c > 0$ *such that for all* N

$$\sigma(1 - P_N) \backslash \cup_{x \in \mathcal{M}_N} \lambda_x \subset (c b_N |\Gamma_N|^{-1}, 2]. \tag{2.7}$$

(iii) For any $x \in \mathcal{M}_N$, *for all* $t > 0$,

$$\mathbb{P}[t_x > t\mathbb{E}\, t_x] = e^{-t(1+o(1))}(1 + o(1)). \tag{2.8}$$

(iv) If ψ_x *denotes the eigenvector of* $1 - P_N$ *corresponding to the eigenvalue* λ_x*, then*

$$\psi_x(y) = \begin{cases} \mathbb{P}[\tau_x^y < \tau_{\mathcal{M}_N(x)}^y](1+o(1)) & \text{if} \quad \mathbb{P}[\tau_x^y < \tau_{\mathcal{M}_N(x)}^y] \geqslant \epsilon_N \\ O(\epsilon_N) & \text{otherwise.} \end{cases} \tag{2.9}$$

Remark. Explicit bounds on the error terms are given [3,6]. In the reversible case we actually have extremely precise control on the $o(1)$ terms in (2.8) in terms of the low-lying eigenvalue cluster, with error terms on the scale determined by the separation from the rest of the spectrum. In the irreversible case, we control the spectrum essentially only in a vicinity of the real axis which replaces (2.7); while this is enough to obtain (2.8), bounds on the errors there are much weaker.

Let us make some additional comments on this theorem. First of all, the identification of what constitutes a metastable exit is crucial, and, in particular, the fact that these processes include the *transition through the 'saddle point'*, guaranteed in our case by the insistence that the process by time t_x has actually arrived in $\mathcal{M}_N(x)$. Without taking this into account, the precise uncertainty principle (ii) could not hold. It is interesting to note that, on the level of this theorem, the difficulties associated with the control of the passage through a saddle are not visible, and that we have the exact formula (2.5) for the mean exit time. Of course the difficulty is hidden in the quantities $T_{x,y}$, whose computation is far from trivial. However, we have shown in [2] that at least in the reversible case, using a variational representation, very precise control of such quantities can be gained in certain settings. Somewhat less precise results can also be obtained in some non-reversible situations [6]. Concerning our estimate of the eigenfunctions, it is easy to see that [3] $\mathbb{P}[\tau_x^y < \tau_{\mathcal{M}_N(x)}^y]$ is either very close to unity or very close to zero, except on a set of points whose invariant measure is extremely small. Therefore, the corresponding right eigenfunctions $\psi_x^r(z) = \mathbb{Q}_N(z)\psi_x(z)$ are essentially proportional to the measure $\mathbb{Q}_N$ conditioned on the local valley corresponding to x (all up to errors of order ϵ_N), i.e. they do indeed represent *metastable measures*, as suggested in [9].

3. Some ideas of the proofs

The first major ingredient of the proofs is a representation formula for the Green function of the transition matrix P_N in terms of certain probabilities. It implies in particular that for any $I \subset \Gamma_N$,

$$\mathbb{E}\, t_I^x = \sum_{y \in \Gamma_N \setminus I \setminus x} \frac{\mathbb{Q}_N(y)}{\mathbb{Q}_N(x)} \mathbb{P}[\tau_x^y < \tau_I^y]\mathbb{P}[\tau_I^x < \tau_x^x] + \frac{1}{\mathbb{P}[\tau_I^x < \tau_x^x]}. \tag{3.1}$$

This formula was first derived for the reversible case in [2]. An apparently independent derivation that also covers the non-reversible case was given recently in [8]. It allows us in particular to prove (2.5) in a rather simple way. However, the realization that this formula actually arises from a representation of the Green function makes it even more useful.

Our analysis of the spectrum of $1 - P_N$ passes through the analysis of the Laplace transforms, $G_{y,J}^x(u) \equiv \mathbb{E}\, e^{ut_y^x} \mathbb{I}_{\tau_y^x < \tau_J^x}$ of transition times of a process that is 'killed' upon arrival in a set $J \subset \Gamma_N$. We write P_N^J for the transition matrix of such a process, and we write λ_J for the smallest eigenvalue of $(1 - P^J)$. It then turns out that all eigenvalues of $(1 - P_N)$ below λ_J can be characterized as follows. Set $u(\lambda) \equiv -\ln(1-\lambda)$. Define the $|J| \times |J|$ matrix $\mathcal{G}_J(u)$ whose elements are

$$\delta_{m',m} - G_{m,J}^{m'}(u) \quad m', m \in J. \tag{3.2}$$

Then λ is an eigenvalue of $(1 - P_N)$ below λ_J if and only if

$$\det G_J(u(\lambda)) = 0. \tag{3.3}$$

This equation is rather easy to understand if $|J| = 1$. In this case, (3.3) becomes simply $G^m_m(u(\lambda)) = 1$. By a simple renewal argument, one sees that $G^m_x(u) = \frac{G^m_{x,m}(u)}{1-G^m_{m,x}(u)}$. Therefore, $u(\lambda)$ defined by (3.3) is the first value at which $\sup_{x\in G_N} G^m_x(u) = +\infty$. The general formula (3.2) is somewhat less intuitive. Basically, one makes an ansatz for the eigenfunctions of $(1 - P_N)$ in terms of the Laplace transforms of the form

$$\Psi(x) = \sum_{m\in J} \phi_m G^x_{m,J}(u). \tag{3.4}$$

One then finds that condition (3.3) is sufficient for the ansatz to yield eigenfunctions with $u = u(\lambda)$. Moreover, one can show that if λ is an eigenvalue then the eigenfunctions can be represented in this form and (3.3) must be satisfied.

To complete the proof one needs good control over the Laplace transforms; this is partly provided again by the representation of the Green function, complemented by lower bounds on eigenvalues λ_J obtained from a Donsker–Varadhan [5] argument. The actual proofs are rather involved and must be left to the longer publications [3, 6]. Let us finally mention that the good control over the spectrum of $(1 - P_N)$ allows a very good control of the analytic properties of the Laplace transforms, which allow in turn the sharp estimates on the probability distribution of metastable transition times stated under (iv).

References

[1] Ben Arous G, Bovier A and Gayrard V 2000 Aging in the random energy model under Glauber dynamics, in preparation

[2] Bovier A, Eckhoff M, Gayrard V and Klein M 2000 Metastability in stochastic dynamics of disordered mean-field models *Prob. Theor. Rel. Fields* at press
(Bovier A, Eckhoff M, Gayrard V and Klein M 2000 *WIAS Preprint* 452)

[3] Bovier A, Eckhoff M, Gayrard V and Klein M 2000 Metastability and low-lying spectra in reversible Markov chains *Commun. Math. Phys.* submitted
(Bovier A, Eckhoff M, Gayrard V and Klein M 2000 *WIAS Preprint* 601)

[4] Biroli G and Kurchan J 2000 Metastable states in glassy systems, webpage http://www.xxx.lanl.gov/cond-mat/0005499

[5] Donsker M D and Varadhan S R S 1976 On the principal eigenvalue of second-order elliptic differential operators *Commun. Pure Appl. Math.* **29** 595–621

[6] Eckhoff M and Klein M 2000 Metastability and low lying spectra in non-reversible Markov chains, in preparation

[7] Freidlin M I and Wentzell A D 1984 *Random Perturbations of Dynamical Systems* (Berlin: Springer)

[8] Gaveau B and Moreau M 2000 Metastable relaxation times and absorbtion probabilities for multidimensional stochastic systems *J. Phys. A: Math. Gen.* **33** 4837–50

[9] Gaveau B and Schulman L S 1998 Theory of nonequilibrium first-order phase transitions for stochastic dynamics *J. Math. Phys.* **39** 1517–33

[10] Mathieu P 1995 Spectra, exit times and long times asymptotics in the zero white noise limit *Stoch. Stoch. Rep.* **55** 1–20

[11] Olivieri E and Scoppola E 1995 Markov chains with exponentially small transition probabilities: first exit problem from a general domain—I. The reversible case *J. Stat. Phys.* **79** 613

[12] Olivieri E and Scoppola E 1996 Markov chains with exponentially small transition probabilities: first exit problem from a general domain—II. The general case *J. Stat. Phys.* **84** 987–1041

[13] Schütte Ch 1999 Conformational dynamics: modelling, theory, algorithm, and application to bio-molecules *ZIB-Berlin Preprint* SC 99–18

[14] Schütte Ch, Fischer A, Huisinga W and Deuflhard P 1999 A direct approach to conformational dynamics based on hybrid Monte Carlo *J. Comput. Phys.* **151** 146–68

[15] Scoppola E 1995 Renormalization and graph methods for Markov chains *Advances in Dynamical Systems and Quantum Physics (Capri, 1993)* (Singapore: World Scientific) pp 260–81

[16] Wentzell A D 1972 On the asymptotic behaviour of the greatest eigenvalue of a second order elliptic differential operator with a small parameter in the higher derivatives *Sov. Math.–Dokl.* **13** 13–7