Invariant measures and Stochastic Discretisations of Dynamical Systems

Rua Murray

Statistical Laboratory, University of Cambridge,
16 Mill Lane, Cambridge CB2 1SB, UK
e-mail: r.murray@statslab.cam.ac.uk

KEYWORDS: Spatial discretisation, Invariant measures, Markov approximation

ABSTRACT

Let $T : X \to X$ be a given map, preserving a probability measure μ. One of the fundamental questions about the ergodic theory of *spatial discretisations* of (T, X) is whether the corresponding transformations of the discrete spaces preserve measures which are "close" to μ. By introducing *stochastic discretisation* (certain Markov Chain approximations) we are able to preserve exactly any invariant μ. However, the necessary discretisations may be difficult to simulate in practice, so we concentrate instead on an easily implementable version. We describe how modelling the discrete transformation by this particular Markov Chain can lead to rigorous bounds on the error for about approximations of absolutely continuous invariant measures.

1 SPATIAL DISCRETISATION AND INVARIANT MEASURES

The ergodic theory of dynamical systems has developed into a very mature subject. For many classes of systems, the set of invariant measures is very well understood, the hierarchy of possible mixing conditions has been well studied, and many fundamental dynamical properties—such as the existence of Lyapounov Exponents—have strong mathematical foundations. For example, see [9]. However, one of the enduring barriers to the rigorous application of these theoretical developments to "real" systems is the (sometimes immense) difficulty in obtaining concrete specifications of appropriate invariant measures. Here, we describe a rigorous numerical procedure for approximating important invariant measures, and explain how it has a natural interpretation as a *stochastic (spatial) discretisation.*

DEFINITION: (INVARIANT MEASURE FOR A DYNAMICAL SYSTEM) Let X be a compact Riemannian manifold, with metric $\rho(\cdot,\cdot)$ and normalised Lebesgue measure $m(\cdot)$. Let $T : X \to X$ be a continuous transformation. Throughout, the pair (X, T) will be our *dynamical system.* A probability measure $\mu(\cdot)$ is called an *invariant measure* if

$$\mu\left(T^{-1}A\right) = \mu(A)$$

for every measurable subset $A \subset X$. $\square$

Because we are interested in using a computer to calculate invariant measures, it is necessary to deal with the computer induced phenomenon of *spatial discretisation*. Indeed, we will be interested in an increasingly fine sequence of such discretisations.

DEFINITION: (SPATIAL DISCRETISATION) Let $X_n \subset X$ be a finite set, consisting of $|X_n|$ distinct points $\{x^{(n),i}\}_{i=1}^{|X_n|}$. Moreover, suppose that there exists a *partition* $\xi^{(n)}$ of X into $|X_n|$ measurable subsets $X^{(n),i}$ such that for each $i = 1 \ldots |X_n|$

$$\mathrm{diam}(X^{(n),i}) = \sup_{y,z \in X^{(n),i}} \rho(y,z) \leq \frac{1}{n}$$

and

$$X_n \cap X^{(n),i} = \{x^{(n),i}\}.$$

The pair $(X_n, \xi^{(n)})$ will be called a $\frac{1}{n}$–*spatial discretisation* of X. Finally, put

$$T_n(x^{(n),i}) = x^{(n),j}$$

where j is the unique index such that $T(x^{(n),i}) \in X^{(n),j}$. Then the pair (X_n, T_n) is a (deterministic) *spatial discretisation* of the dynamical system (X, T). □

Luckily, probability measures are just as easy to discretise. If μ is a probability measure on X, then

$$\mu^{(n)} = \sum_{i=1}^{|X_n|} \mu(X^{(n),i}) \delta_{x^{(n),i}}$$

is a probability measure on X_n where δ_x is the point measure

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in X \setminus A. \end{cases}$$

Conversely, any probability measure $\nu^{(n)}$ on X_n is a sum of point measures, and because X_n can be naturally embedded in X, $\nu^{(n)}$ is also a probability measure on X. Moreover, it easy to check that any *invariant probability measure* for (T_n, X_n) is a sum of measures which are uniformly concentrated on periodic orbits of the discrete transformation.

PROPOSITION 1: (WEAK CONVERGENCE TO INVARIANT MEASURES) *Let (X_n, T_n) be a sequence of $\frac{1}{n}$–spatial discretisations of (T, X), and for each n let $\mu^{(n)}$ be an invariant probability measure for T_n. Then any weak–$*$ limit point of the sequence $\{\mu^{(n)}\}$ is an invariant probability measure for T.*

REMARK: Because the weak–$*$ topology of probability measures on X is compact, each sequence $\{\mu^{(n)}\}$ has at least one limit point. The Proposition can be proved as an easy application of a certain metric for the weak–$*$ topology for probability measures on X [11]. For related results, see [3]. □

Proposition 1 says that invariant measures are, in some sense, robust to spatial discretisation. (We assume that finding an invariant measure for (X_n, T_n)—being a finite problem—is easy to do, at least in principle). However, it is obviously impossible to compute invariant measures for an infinite sequence of spatial discretisations. Moreover, we have neither information about *which* invariant measures are approached asymptotically, nor the rate of convergence. The *stochastic discretisation* based methods described below have none of these drawbacks.

2 STOCHASTIC DISCRETISATION AND A MEASURE PRESERVING MARKOV MODEL

Above, the emphasis in the discussion of the spatial discretisation of X is on the collections of points X_n. Henceforth, it is preferable to think of the discretisation in terms of the *partitions* $\xi^{(n)}$. Then, the definition of T_n may be rewritten to say

$$T_n(X^{(n),i}) = X^{(n),j}$$

where j is the unique index such that $T(x^{(n),i}) \in X^{(n),j}$. By regarding the discretisation as a collection of subsets, rather than a collection of points, it seems a little unnatural to select $T_n(X^{(n),i})$ based only on the dynamical image of the privileged point $x^{(n),i}$. Indeed, it would be equally reasonable to choose $T_n(X^{(n),i})$ to be *any* of the subsets $X^{(n),j'}$ which contain the image of a point from $X^{(n),i}$. This motivates the following definition:

DEFINITION: (STOCHASTIC DISCRETISATION) Let $\xi^{(n)}$ be the partition corresponding to a $\frac{1}{n}$–spatial discretisation of X, and let $P_n(\cdot; i)$ be a family of probability measures on $\xi^{(n)}$ indexed by $\{i : X^{(n),i} \in \xi^{(n)}\}$. The family $\{P_n(\cdot; i)\}$ defines a Markov Chain $\{Z_k\}_{k\geq 0}$ on $\xi^{(n)}$ by the formula

$$\mathbb{P}\left(Z_{k+1} = X^{(n),j} | Z_k = X^{(n),i}\right) = P_n\left(X^{(n),j}; i\right).$$

If the family $\{P_n(\cdot; i)\}$ has the property that

$$P_n(X^{(n),j}; i) > 0 \quad \Rightarrow \quad T(X^{(n),i}) \cap X^{(n),j} \neq \emptyset$$

then the Markov Chain is called a $\frac{1}{n}$*–stochastic discretisation* of (X, T). A measure $\mu^{(n)}$ will be said to be an *invariant measure* for a stochastic discretisation if and only if it is an invariant measure for the Markov Chain. That is,

$$\mu^{(n)}\left(X^{(n),j}\right) = \sum_{i=1}^{|X_n|} \mu^{(n)}\left(X^{(n),i}\right) P_n\left(X^{(n),j}; i\right). \quad \square$$

EXAMPLE: Deterministic spatial discretisations are a special case of stochastic discretisations. Then,

$$P_n\left(X^{(n),j}; i\right) = \begin{cases} 1 & \text{if } T(x^{(n),i}) \in X^{(n),j}, \\ 0 & \text{otherwise.} \end{cases}$$

It is an easy exercise to check that any invariant measure for (X_n, T_n) is an invariant measure for the corresponding stochastic discretisation. $\square$

REMARK: With a suitable choice of partition, the class of random perturbations described by Diamond *et. al.* [2] can be regarded as stochastic discretisations. For more background, see the references contained therein. $\square$.

From [11]:

PROPOSITION 2: (INVARIANT MEASURES BY STOCHASTIC DISCRETISATION [11]) *Let* $(\xi^{(n)}, \{P_n(\cdot; \cdot)\})$ *be a sequence of* $\frac{1}{n}$*–stochastic discretisations of* (X, T) *and let* $\{\mu^{(n)}\}$ *be a corresponding sequence of invariant measures. Regarding each* $\mu^{(n)}$ *as a probability measure on* X*, any weak–$*$ limit point of* $\{\mu^{(n)}\}$ *is an invariant measure for* T*.*

REMARK: Notice that any sample path of a $\frac{1}{n}$–stochastic discretisation can be embedded as a $\frac{1}{n}$–pseudo–orbit of T: let $\{Z_k\}_{k\geq 0}$ be a sample path of the Markov Chain on $\xi^{(n)}$, and for each k let z_k be any point in $Z_k \cap T^{-1}(Z_{k+1})$ (the intersection is non–empty by the definition of stochastic discretisation). Then, both

$$T(z_k), z_{k+1} \in Z_{k+1},$$

so that $\rho(T(z_k), z_{k+1}) \leq \frac{1}{n}$, and $\{z_k\}_{k\geq 0}$ is a $\frac{1}{n}$–pseudo–orbit, as claimed. Because the most one can say in general about a deterministic spatial discretisation is that it's trajectories are

$\frac{1}{n}$–pseudo–orbits, there is no loss of dynamical accuracy in admitting the *random* trajectories of a stochastic discretisation. $\Box$.

Now, suppose that $T : X \to X$ preserves a probability measure μ. For each n let $(X_n, \xi^{(n)})$ be a $\frac{1}{n}$–spatial discretisation of X, and for each $i, j = 1 \ldots |X_n|$ put

$$P_n(X^{(n),j}; i) = \begin{cases} \frac{\mu\left(T^{-1}\left(X^{(n),j}\right) \cap X^{(n),i}\right)}{\mu\left(X^{(n),i}\right)} & \text{if } \mu(X^{(n),i}) > 0, \\ 1 & \text{if } \mu(X^{(n),i}) = 0 \text{ and } i = j, \\ 0 & \text{if } \mu(X^{(n),i}) = 0 \text{ and } i \neq j. \end{cases} \qquad (*)$$

Then, it is an easy exercise to show that:

PROPOSITION 3: *The stochastic discretisations defined by (*) have μ as an invariant measure.*

Therefore, we can choose arbitrarily fine stochastic discretisations to reproduce *any* given invariant probability measure. However, Proposition 3 is primarily of *illustrative* rather than *computational* value. This is because exact knowledge of the necessary transition probabilities would imply knowledge of the required invariant measure, thereby obviating the need to compute it.

3 A COMPUTABLE MARKOV MODEL

Of particular interest are invariant measures which are *absolutely continuous* with respect to the Lebesgue measure m. For the remainder of the paper, we discuss a particular choice of stochastic discretisation: Ulam's method [12]. The method is reasonably well known, well studied [8, 5, 10] (and many others), and can be written down as a stochastic discretisation. Let $(X_n, \xi^{(n)})$ be a $\frac{1}{n}$–spatial discretisation of X, and for each $i, j = 1 \ldots |X_n|$ put

$$P_{ij}^{(n)} = P_n(X^{(n),j}; i) = \frac{m\left(T^{-1}(X^{(n),j}) \cap X^{(n),i}\right)}{m\left(X^{(n),i}\right)}. \qquad (**)$$

Then $P^{(n)} = (P_{ij}^{(n)})$ is a *stochastic matrix* and the transition probabilities define a stochastic discretisation of (X, T). Ulam's method consists in finding an invariant measure $p^{(n)} = (p_i^{(n)})_{i=1}^{|X_n|}$ for the corresponding Markov Chain, and letting $\mu^{(n)}$, a probability measure on X, be defined setwise by

$$\mu^{(n)}(A) = \sum_{i=1}^{|X_n|} p_i^{(n)} \frac{m(A \cap X^{(n),i})}{m(X^{(n),i})}$$

for every measurable set $A \subset X$. By Proposition 2, any weak–$*$ limit point of the sequence $\{\mu^{(n)}\}$ generated by Ulam's method is an invariant measure for T.

Superficially, Ulam's method does not look any better than the general scheme for stochastic discretisation described above. However, it is known that for piecewise expanding transformations of the interval [8] and rectangles in $\mathbb{R}^d$ [4], the sequence of measures given by Ulam's method converges strongly to an absolutely continuous invariant measure for the given transformation. Moreover, in some situations it is possible to get strong *quantitative bounds* on the accuracy of the measure approximations. We describe these in the next section.

REMARK: It is actually rather easy to calculate the invariant measures for the stochastic discretisations corresponding to Ulam's method. As mentioned above, for each n this consists in finding the probability vector $p^{(n)} = (p_i^{(n)})_{i=1}^{|X_n|}$. Recalling the definition of an invariant measure for a stochastic discretisation, this implies

$$p_j^{(n)} = \sum_{i=1}^{|X_n|} p_i^{(n)} P_{ij}^{(n)} \quad \forall j = 1 \dots |X_n|,$$

where $P^{(n)}$ is the stochastic matrix defined in (**). Therefore, to get $p^{(n)}$, one needs only to calculate the entries of the matrix $P^{(n)}$, and numerically determine a left eigenvector. Alternatively, the Ergodic Theorem for finite state Markov Chains implies that the frequency of visits to each state of a typical sample path of the corresponding stochastic discretisation will converge to an ergodic component of $p^{(n)}$ as the length of the sample path tends to infinity [10]. □

4 SOME PRECISE APPROXIMATION RESULTS

Having described Ulam's method as a stochastic discretisation, we now report some of our results about the accuracy of the Ulam approximations.

Throughout, $(X_n, \xi^{(n)})$ will denote a $\frac{1}{n}$–spatial discretisation, and $p^{(n)}$ will denote an invariant probability measure of the stochastic discretisation corresponding to Ulam's method. Let μ denote the (unique) absolutely continuous invariant measure (acim) for T (which always exists under the hypotheses we impose below). We now describe several classes of transformations.

($\mathcal{T}1$) Let $X = S^1$ be the unit circle, and let T be an expanding $\mathcal{C}^2$ circle map. That is, there exist constants λ and K such that $|T'| \geq \lambda > 2$ and $|T''|/|T'| \leq K < \infty$.

($\mathcal{T}2$) Let $X = [0,1]$ and let T be a piecewise continuous map, where every monotonicity branch of T maps onto $[0,1]$. Suppose there exist constants λ, s such that $|T'| > \lambda$ and $|T''|/|T'|^2 \leq s$. This class includes the well known continued fraction transformation.

($\mathcal{T}3$) Let $X = [0,1]$, let $\beta > 2$ and let $T : x \mapsto \beta x \pmod 1$.

($\mathcal{T}4$) Let $X = [0,1]^d$, and let T be an expanding $\mathcal{C}^2$ transformation satisfying the following properties: there exist constants $\lambda > 1$, $s < \infty$ and a subset $A \subset X$ with $m(A) > 0$ such that $\|DT\| > \lambda$, $\|d(\log|\det(D(T^{-1}))|)\| \leq s$ and that the image of every monotonicity branch of T covers the set A. Here, $\|\cdot\|$ is the usual Euclidean vector (and matrix) norm, DT denotes the Jacobian matrix of T, and d the gradient. This class includes all piecewise–onto $\mathcal{C}^2$ expanding transformations, and many of the transformations corresponding to multi–dimensional continued fraction algorithms [1].

THEOREM: *For transformations in each of the classes ($\mathcal{T}1$)–($\mathcal{T}4$), there exist constants $C_\mathcal{T}$ depending only on the constants defining the relevant class, such that for every n,*

$$\sum_{i=1}^{|X_n|} \left| p_i^{(n)} - \mu(X^{(n),i}) \right| \leq C_\mathcal{T} \frac{\log n}{n}.$$

In each case, the constant $C_\mathcal{T}$ has an explicit formula.

REMARK: For transformations in ($\mathcal{T}1$), the result is proved in [6]. For ($\mathcal{T}2$) and ($\mathcal{T}3$) see [10], the proof of ($\mathcal{T}4$) is in preparation. The proofs of the Theorem for the various classes of transformations involve careful estimates of the *rates of mixing* of the stochastic discretisations corresponding to Ulam's method. The fundamental idea is that these stochastic discretisations, with transition probabilities constructed from the Lebesgue measure m, are small perturbations of the stochastic discretisations described in Section 2, where the transition probabilities are chosen according to the acim μ. The *rate of mixing* estimates allow the effects of the perturbation to be bounded. $\square$

The results summarised in the Theorem are new. Most previous rigorous work on approximating invariant measures has focussed on analysing the convergence of the scheme in the limit [8, 5, 4]. Keller [7] demonstrated that an $O(\log n/n)$ rate exists, but did not give any constructive bounds on the constants in the $O(\cdot)$ notation. The results from [6, 10] that we mention above differ fundamentally in character because of the explicit quantitative bounds on the various constants. In [6] and [10], the approach is illustated with a variety of examples where the theorems give rigorous bounds on the accuracy of the numerically calculated invariant measure approximations.

REFERENCES

[1] V Baladi and A Nogueira. *Lyapounov exponents for non–classical multi–dimensional continued fraction algorithms.* Nonlinearity, 9:1529–1546, 1996.

[2] P Diamond, P Kloeden, and A Pokrovskii. *An invariant measure arising from a computer simulation of a chaotic dynamical system.* Journal of nonlinear science, 4(1):59–68, 1994.

[3] P Diamond, P Kloeden, and A Pokrovskii. *Analysis of an algorithm for computing invariant measures.* Nonlinear Analysis, 24(3):323–336, 1995.

[4] Jiu Ding and Aihui Zhou. *Finite approximations of Frobenius–Perron operators. A solution of Ulam's conjecture to multi–dimensional transformations.* Physica D, 92:61–68, 1996.

[5] G Froyland. *Finite approximation of Sinai–Bowen–Ruelle measures for Anosov systems in two dimensions.* Random & Computational Dynamics, 3(4):251–264, 1995.

[6] M S Keane, R Murray, and L-S Young. *Computing invariant measures for expanding circle maps.* Submitted for publication.

[7] G Keller. *Stochastic stability in some chaotic dynamical systems.* Monatsh. Mathematik, 94:313–333, 1982.

[8] T-Y Li. *Finite approximation for the Perron–Frobenius operator. A solution to Ulam's conjecture.* Journal of approximation theory, 17:177–186, 1976.

[9] R Mañé. *Ergodic Theory and Differentiable Dynamics*. Springer, 1983.

[10] R Murray. *Approximation error for invariant density calculations.* Submitted for publication.

[11] R Murray. *Weak convergence and the approximation of invariant measures.* In preparation.

[12] S Ulam. *A collection of mathematical problems.* Interscience Publishers, 1960.