Kuzmin, coupling, cones, and exponential mixing

Roland Zweimüller

Institut für Mathematik, Universität Salzburg
Hellbrunnerstraße 34, A-5020 Salzburg, Austria
roland.zweimueller@sbg.ac.at

20 May 2002

Abstract

We study a class of fibred systems with good distortion properties (Gibbs-Markov maps), including Folklore maps as well as multidimensional continued fraction algorithms like Jacobi-Perron. Using an elementary coupling scheme based on regularity we give an easy proof of an exponential uniform convergence (or "Kuzmin type") theorem for the iterates of the transfer operator. This approach is then shown to be equivalent to the cone contraction method.

AMS subject classification: 28D05, 37A25, 37C30, 11K50.

1 Introduction

Perhaps the first instance of a question related to the transfer operator of a dynamical system is to be found in Gauss' letter [Ga] to Laplace, in which (interpreted in modern terms) he announced the invariant density for the continued fraction transformation, claimed that the latter was mixing, and asked for the speed of convergence to the invariant measure. Gauss classified the question on continued fractions he was interested in as one belonging to probability theory, thus anticipating the viewpoint ergodic theory takes today. Kuzmin [Ku] was the first to (partially) answer Gauss' question by giving (stretched exponential) bounds on the rate of convergence, and shortly after that Lévy [Le] showed that convergence is in fact exponentially fast.

In the context of metric number theory and multidimensional continued fractions, results of this type are still referred to as Kuzmin type theorems. Like in the case of one-dimensional systems, techniques of spectral theory can be used to prove exponential convergence, see e.g. [Me], [Br], [Aa], [Be] and [S3] for versions which apply to the multidimensional Jacobi-Perron algorithm. Still there is some persistent interest in more elementary arguments, and a convergence theorem derived by Kuzmin's original approach has only recently been published in [S2], resulting in stretched exponential bounds on the rate, while no easy proof of exponential convergence for this type of systems seems to be available. The purpose of the present note is to point out that it is well possible to prove exponential uniform convergence by a slight variant of Kuzmin's approach which can be interpreted as a simple and completely elementary coupling scheme.

Basically the idea for coupling densities is to start inside some convenient family of (e.g. smooth) initial densities, to transport mass by iterating the transfer operator, and to remove suitable multiples of the invariant density on the way in a manner which allows us to control the mass that remains. The proportion of mass that can be coupled at each step depends on the smoothness of the remaining density *compared to its total mass*. Therefore, employing such a scheme based on control of the derivative alone (as in Kuzmin's original method), one encounters a problem as the remaining mass decreases, which is one way to understand why that approach does not lead to an exponential convergence result. Below we follow the same coupling idea, a priori basing it on control of the *regularity* of densities, i.e. the smoothness of their *normalized* versions, which enables us to show that indeed some fixed proportion of mass can be coupled at each step, thus proving exponential convergence. Finally we discuss the intimate relation of this coupling approach to the cone method (cf. [Ba], [Liv]).

2 Framework and statement of the result

Throughout (X,d) will be a compact metric space with Borel σ-field $\mathcal{B}$, and λ is some (reference) probability measure on $\mathcal{B}$. We shall simply write $\lambda(u)$ for $\int_X u\,d\lambda$. If $\lambda(B)>0$, we let $\lambda_B(A):=\lambda(B)^{-1}\lambda(A\cap B)$. For convenience (and w.l.o.g.) we assume that $diam(X)\leq 1$.

Definition 1 *A realvalued function u on $B\subseteq X$ will be called* admissible on B *if u is Lipschitz on B and $\inf_B u>0$ or, equivalently, if $u>0$ and there is some constant $r>0$ such that*

$$\frac{u(x)}{u(y)}\leq 1+r\cdot d(x,y)\qquad \textit{for } x,y\in B. \tag{1}$$

In this case we define the regularity of u on B *as $R_B(u):=\inf\{r>0: r$ satisfies (1)$\}$. Then $r:=R_B(u)$ itself satisfies (1). The constant function $u=0$ on B will also be regarded admissible with $R_B(0):=0$. If β is a collection of subsets of X, a function u on X is* β-admissible *if for each $B\in\beta$ the restriction $u\mid_B$ is admissible on B, and if the* β-regularity of u, *defined as $R_\beta(u):=\max\{R_B(u): B\in\beta\}$, is finite.*

Remark 1 *Let $Lip_B(u)$ denote the Lipschitz constant of u on B, then clearly $R_B(u)\leq Lip_B(u)/\inf_B u$. If $X\subset\mathbb{R}^m$, d is the Euclidean metric, B is convex, and $u\in\mathcal{C}^1(B)$ with $\|Du\|/u\leq s$, then $R_B(u)\leq e^s$ (with obvious generalizations to sets which are boundedly path-connected). Our notion of regularity thus is closely related to the concept used in [PY] or [Zw] (but does not precisely agree with it).*

Remark 2 **a)** *If u is admissible on B, then $\frac{u(x)}{u(y)}\leq 1+R_B(u)$ for $x,y\in B$, and therefore*

$$\frac{\lambda_B(u)}{1+R_B(u)}\leq u\leq(1+R_B(u))\,\lambda_B(u).$$

b) *Recall that any Lipschitz function on B has a (unique) Lipschitz extension to $cl(B)$. We shall tacitly identify these two functions in the sequel.*

Definition 2 *We consider* nonsingular piecewise invertible (or fibred) systems *on X, i.e. triples (X,T,ξ) where ξ is a partition (mod λ) of X into a finite or countable number of open subsets (the* cylinders of rank one*), and $T: X\to X$ is a surjective (mod λ) map such that each of the restrictions $T\mid_Z$, $Z\in\xi$, is a nonsingular (w.r.t. λ) homeomorphism of Z onto TZ.*

Given such a system (X, T, ξ), we let ξ_n denote the family of *cylinders of rank* n, that is, the nonempty sets of the form $Z = [Z_0, \dots, Z_{n-1}] := \bigcap_{i=0}^{n-1} T^{-i} Z_i$ with $Z_i \in \xi$. We let f_Z be the inverse of the branch $T^n \mid_Z$. Then $f_Z : T^n Z \to Z$ has a Radon-Nikodym derivative $\omega_Z := d(\lambda \circ f_Z)/d\lambda$ on $T^n Z$. $\xi(x)$ is the member of ξ containing x. The fundamental partition ξ respectively the system (X, T, ξ) are said to be *Markov* if $TZ \cap Z' \neq \varnothing$ implies $Z' \subseteq TZ$ whenever $Z, Z' \in \xi$. In this case there is an *image partition* β (i.e. a coarsest partition with respect to which each TZ, $Z \in \xi$ is measurable) which is refined by ξ. We shall consider Markov systems (X, T, ξ) which satisfy the *finite image condition*, i.e.

$$T\xi = \{TZ : Z \in \xi\} \text{ is finite.} \tag{$\mathbf{F}$}$$

In this case the image partition β is finite, too. Moreover, the transition structure of the system should be *aperiodic on* β in the sense that

$$\inf_{B_0, B_1 \in \beta} \lambda(B_0 \cap T^{-n} B_1) > 0 \text{ for some (and hence all sufficiently large) } n. \tag{$\mathbf{AP}$}$$

Also, the systems will have a *uniformly expanding* iterate, i.e. there are $N \geq 1$ and $\rho = \rho(T^N) \in (0, 1)$ for which

$$d(f_Z(x), f_Z(y)) \leq \rho \cdot d(x, y) \quad \text{ for } x, y \in Z \in \xi_N. \tag{$\mathbf{U}$}$$

Finally, to obtain enough control of distortion, we will suppose that (X, T, ξ) satisfies the following abstract version of the folklore "*Adler's condition*" requiring that the ω_Z have β-admissible versions for which

$$A = A(T) := \sup_{Z \in \xi} R_{TZ}(\omega_Z) < \infty. \tag{$\mathbf{A}$}$$

Remark 3 **a)** *Standard calculations show that* **(U)** *and* **(A)** *together imply*

$$A_\infty = A_\infty(T) := \sup_{n \geq 1} \sup_{Z \in \xi_n} R_{T^n Z}(\omega_Z) < \infty, \tag{$\mathbf{G}$}$$

which is (equivalent to) the Gibbs property *of [Aa]. Consequently,* (X, T, ξ) *also satisfies* Rényi's condition*: There is some* $C_{\mathcal{R}} \geq 1$ *such that for all* $n \geq 1$ *and* $Z \in \xi_n$ *we have*

$$\sup_{T^n Z} \omega_Z \leq C_{\mathcal{R}} \cdot \inf_{T^n Z} \omega_Z \ . \tag{$\mathbf{R}$}$$

b) *In any dimension* $m \geq 1$, *if* $X \subset \mathbb{R}^m$ *and* ξ *is a finite collection of sets as in Remark 1 then* **(A)** *holds automatically if each* f_Z, $Z \in \xi$, *has a* $\mathcal{C}^2$ *extension to the closure of* TZ. **c)** *If* d_1 *and* d_2 *are Lipschitz equivelent metrics, then condition* (**A**) *either holds for both or for none of them.*

We are interested in the *transfer (Perron-Frobenius or dual) operator* $\mathbf{P}$ *of* T *with respect to* λ, which is characterized by $\int_X \mathbf{P}u \cdot f \, d\lambda = \int_X u \cdot f \circ T \, d\lambda$ for all $u \in L_1(\lambda)$ and $f \in L_\infty(\lambda)$. Its nth power $\mathbf{P}^n$ is the transfer operator of T^n and has an explicit representation as

$$\mathbf{P}^n u = \sum_{Z \in \xi_n} 1_{T^n Z} (u \circ f_Z) \cdot \omega_Z \ , \quad n \geq 1. \tag{2}$$

The asymptotic behaviour of the powers $\mathbf{P}^n$ of the transfer operator is of central interest for the finer ergodic and probabilistic properties of the system. We are going to give an elementary proof of the following exponential uniform convergence theorem for admissible initial densities.

Theorem 1 **(Exponential convergence to the invariant density)** *Let (X,T,ξ) be a Markov map satisfying* **(F)***,* **(AP)***,* **(U)***, and* **(A)***. Then T has a unique invariant probability density, which has a β-admissible version $h>0$, and there are constants $q\in(0,1)$ and $H,K\in(0,\infty)$ such that for any $u>0$ admissible on X we have*

$$\|\mathbf{P}^n u-\lambda(u)\,h\|_\infty \le K\,(H+R_\beta(u))\;\lambda(u)\cdot q^n \qquad \textit{for } n\ge 0.$$

3 Examples

We give a few examples illustrating the scope of the result.

Example 1 **(Folklore interval maps)** *For $X=[0,1]$, d the Euclidean metric, and λ Lebesgue measure, we recover the well known Folklore class of interval maps satisfying the classical (Euclidean) Adler condition.*

Example 2 **(Hölder continuous derivatives)** *Our setup is flexible enough to include interval maps violating the classical Adler condition but satisfying a related Hölder condition: Assume that instead of (*__A__*) the system (X,T,ξ) only satisfies $\omega_Z(x)/\omega_Z(y)\le 1+A\,d(x,y)^\gamma$ for $Z\in\xi$ and $x,y\in TZ$ with some $A>0$ and $\gamma\in(0,1)$. Then the assomptions of Theorem 1 are satisfied w.r.t. the metric $\delta(x,y):=d(x,y)^\gamma$ on X. (And any u admissible for d is also admissible for δ.)*

Example 3 **(The m-dimensional Jacobi-Perron algorithm)** *We refer to [S1], [S3], or [Br] for some basic information on this multidimensional continued fraction algorithm which satisfies all the assumptions of our theorem. The result thus applies without difficulties, providing us with an elementary proof of an exponential "Kuzmin theorem" for the Jacobi-Perron algorithm.*

Example 4 **(Finite-range systems. The class studied in [Be])** *The piecewise invertible systems (X,T,ξ) considered in [S2] and [Be] need not be Markov in the first place, but are assumed to have* finite range structure*, cf. condition (B) there, i.e.*

there is a finite collection $\mathcal{U}$ of image sets s.t. for any $n\ge 1$ and $Z\in\xi_n$ we have $T^nZ\in\mathcal{U}$ (and each $U\in\mathcal{U}$ should occur in this way). (**FRS**)

*However, this condition implies that they can essentially be treated as Markov systems, and we recall two ways of doing so. First, finite range structure yields a Markov partition: Let (X,T,ξ) satisfy (*__FRS__*) and β denote the finite partition generated by $\mathcal{U}$. Then it is easy to see (compare [Yu]) that $\xi':=\xi\vee\beta$ is a Markov partition refining ξ so that the Markov system (X,T,ξ') satisfies (*__F__*) with the same image partition β. In general, the geometry of this refined partition may be worse, but in many of the concrete examples, this is not the case. Second, a way to generally reduce (*__FRS__*)-systems to Markov situations is to systematically separate the different image sets by passing to the (finite!)* canonical Markov extensions (C.M.E.s) *of the system, built from image sets as in [Ke] or [Zw].*

*Let (X,T,ξ) be a fibered system satisfying the assumptions of [Be]. Then this is a nonsingular piecewise invertible system in our sense, satisfying (*__FRS__*). Conditions (A) and (F') there imply (*__U__*), and in the presence of the Renyi property (*__R__*) (condition (C) there), (*__A__*) is equivalent to condition (E) of [Be]. Furthermore, condition (D) there implies (*__AP__*), see the remark on p.559 of [S2]. Therefore Theorem 1 applies to the C.M.E. of (X,T,ξ) (and so does the spectral theory of Gibbs-Markov maps as developed in [Aa]). The exponential convergence result immediately carries over to the original system which is a bounded-to-one factor of its extension.*

4 Proof of the theorem

Throughout this section (X, T, ξ) denotes some fixed Markov system satisfying **(F)**, **(AP)**, **(U)**, and **(A)**. We begin with some convenient simplifications: Since $\rho(T^{kN}) \leq \rho^k$, $k \geq 1$, and $A(T^n) \leq A_\infty(T)$, $n \geq 1$, we can choose $n_0 \geq 1$ in such a way that $\rho(T^{n_0})(1 + A(T^{n_0})) < 1$ and so that it also satisfies the aperiodicity property $\lambda(B_0 \cap T^{-n_0} B_1) > 0$ for $B_i \in \beta$.

Assume that the theorem is proved for T^{n_0}. If $n = kn_0 + l$, $0 \leq l < n_0$, then

$$\|\mathbf{P}^n u - \lambda(u)\, h\|_\infty \leq \frac{1 + A_\infty}{\min_{B\in\beta} \lambda(B)} \cdot \left\|\mathbf{P}^{kn_0} u - \lambda(u)\, h\right\|_\infty,$$

since $\mathbf{P}$ is positive and $\mathbf{P}^l 1 \leq (1 + A_\infty)/\min_{B\in\beta} \lambda(B)$ for any $l \geq 1$. Therefore we may assume w.l.o.g. that

$$N = 1,\ \rho_0 := \rho(1 + A) < 1, \text{ and } \lambda(B_0 \cap T^{-1} B_1) > 0 \text{ for } B_i \in \beta. \tag{3}$$

The first important observation is a Doeblin-Fortet type inequality for β-regularity under the action of the transfer operator (compare [PY]).

Lemma 1 (Regularity and the transfer operator) *Let u be β-admissible, then so is $\mathbf{P}u$, and*

$$R_\beta(\mathbf{P}u) \leq \rho_0 R_\beta(u) + A,$$

where $\rho_0 := \rho(1 + A)$. In particular, $R_\beta(\mathbf{P}u) \leq \max(R_\beta(u), A_0)$, where $A_0 := (1 - \rho_0)^{-1} A$.

Proof. Fix any $B \in \beta$. Since ξ refines β, we have $R_Z(u) \leq R_\beta(u)$ whenever $Z \in \xi$. For $x, y \in B$ therefore

$$\begin{aligned}
\mathbf{P}u(x) &= \sum_{Z\in\xi:TZ\supseteq B} u(f_Z(x)) \cdot w_Z(x) \\
&\leq \sum_{Z\in\xi:TZ\supseteq B} (1 + R_Z(u)\, d(f_Z(x), f_Z(y)))\, u(f_Z(y)) \cdot (1 + A\, d(x, y))\, w_Z(y) \\
&\leq (1 + \rho R_\beta(u)\, d(x, y))\, (1 + A\, d(x, y)) \cdot \mathbf{P}u(y).
\end{aligned}$$

Hence $R_B(\mathbf{P}u) \leq \rho(1 + A) R_\beta(u) + A$. Moreover, for any t, $\rho_0 t + A = \rho_0 t + (1 - \rho_0) A_0 \leq \max(t, A_0)$. ■

Existence of a β-admissible invariant density. Let u be any admissible probability density with $R_\beta(u) \leq A_0$ (e.g. $u := 1_X$). By Lemma 1, $R_\beta(\mathbf{P}^n u) \leq A_0$ and, *a fortiori*, $R_\beta(n^{-1} \sum_{k=0}^{n-1} \mathbf{P}^k u) \leq A_0$ for all $n \geq 0$. The usual Arzela-Ascoli argument (on each $cl(B)$, $B \in \beta$, cf. Remark 2 b), then diagonalizing) provides us with a β-admissible invariant probability density h with $R_\beta(h) \leq A_0$. By (**AP**), $h > 0$ on X. ■

Exponential convergence. Our proof of exponential convergence carries out the coupling idea sketched in the introduction. At each iteration step we are going to remove a certain proportion $p \in (0, 1)$ of mass from a β-admissible function u by subtracting the appropriate multiple of $v = h$. For β-admissible function v and $p > 0$ we define the *coupling operator* $\triangle_{p,v}$ by

$$\triangle_{p,v}(u) := u - p \cdot \frac{\lambda(u)}{\lambda(v)}\, v, \quad u \text{ any } \beta\text{-admissible function on } X.$$

The key to succesful coupling is the following information about the behaviour of regularity under this operation:

Lemma 2 (Regularity and the coupling operator) *Let u, v be β-admissible and assume that $0 < \kappa^{-1} \leq \lambda_B(u)/\lambda(u) \leq \kappa$ for each $B \in \beta$, and that the same estimate holds for v. Let $p > 0$ be so small that $p(1+R_\beta(u))(1+R_\beta(v))\kappa^2 < 1$. Then $w := \triangle_{p,v}(u)$ is positive, β-admissible, $\lambda(w) = (1-p)\lambda(u)$, and*

$$R_\beta(w) \leq R_\beta(u) + p\frac{(1+R_\beta(u))(1+R_\beta(v))(R_\beta(u)+R_\beta(v))\kappa^2}{1-p(1+R_\beta(u))(1+R_\beta(v))\kappa^2} \ .$$

Proof. Fix any $B \in \beta$. For $x, y \in B$ we find (recalling Remark 2 a))

$$\begin{aligned} w(y) &= u(y)\left(1 - p\frac{\lambda(u)}{u(y)}\frac{v(y)}{\lambda(v)}\right) \\ &\geq u(y)\left(1 - p(1+R_\beta(u))(1+R_\beta(v))\frac{\lambda(u)}{\lambda_B(u)}\frac{\lambda_B(v)}{\lambda(v)}\right) > 0, \end{aligned}$$

and

$$\begin{aligned} \frac{w(x)}{w(y)} &= 1 + \frac{\frac{u(x)}{u(y)} - 1 + p\frac{\lambda(u)}{u(y)}\frac{v(x)}{\lambda(v)}\left(\frac{v(y)}{v(x)} - 1\right)}{1 - p\frac{\lambda(u)}{u(y)}\frac{v(x)}{\lambda(v)}} \\ &\leq 1 + \frac{R_B(u) + p(1+R_B(u))(1+R_B(v))\kappa^2 R_B(v)}{1-p(1+R_B(u))(1+R_B(v))\kappa^2} \cdot d(x,y), \end{aligned}$$

which gives the asserted bound. ■

To apply this estimate we shall need some control of the κ appearing there, i.e. of the mass fluctuations between the members of β. This is what the next observation provides us with.

Lemma 3 (Balance of mass between the atoms of β) *For any $r > 0$ there is some $\kappa(r) > 1$ such that for any β-admissible u with $R_\beta(u) \leq r$ and any $B \in \beta$,*

$$\kappa(r)^{-1} \leq \frac{\lambda_B(\mathbf{P}u)}{\lambda(\mathbf{P}u)} \leq \kappa(r).$$

Proof. Trivially, $\lambda_B(\mathbf{P}u) \leq (\min_{B\in\beta}\lambda(B))^{-1}\lambda(\mathbf{P}u)$. On the other hand, $c_\beta := \min_{B_i\in\beta}\lambda(B_0 \cap T^{-1}B_1) > 0$, and given u, there is $B_0 = B_0(u) \in \beta$ for which $\lambda_{B_0}(u) \geq \lambda(u) = \lambda(\mathbf{P}u)$. For arbitrary $B_1 \in \beta$ we therefore have (recalling Remark 2 a) and Lemma 1),

$$\begin{aligned} \lambda_{B_1}(\mathbf{P}u) &\geq (\min_{B\in\beta}\lambda(B))^{-1}\int_{B_o\cap T^{-1}B_1} u\, d\lambda \\ &\geq (\min_{B\in\beta}\lambda(B))^{-1} c_\beta \inf_{B_0} u \\ &\geq \left(\min_{B\in\beta}\lambda(B)\right)^{-1}\frac{c_\beta}{(1+R_\beta(u))}\lambda_{B_0}(u). \qquad ■ \end{aligned}$$

We can now verify that the coupling scheme advertised above works, thus completing the proof of the theorem. Assume that $u =: u_0$ is admissible on X, $\lambda(u) > 0$, and suppose for the moment that $R_\beta(u) \leq r_0 := 2A_0$. We are going to show that there is some $p \in (0,1)$ such that inductively defining

$$\widetilde{u}_n := \mathbf{P}u_{n-1}, \quad u_n := \triangle_{p,h}(\widetilde{u}_n), \quad \text{for } n \geq 1,$$

gives a sequence (u_n) of β-admissible positive functions on X with

$$R_\beta(\widetilde{u}_k),\, R_\beta(u_k) \leq r_0 \ , \tag{4}$$

for all $k \geq 1$. If this is the case, we have $\mathbf{P}^n u - \lambda(u)\, h = u_n - \lambda(u_n) \cdot h$. Since $\lambda(u_n) = (1-p)^n \lambda(u)$, and (using Lemma 3 and Remark 2 a)) $\sup u_n \leq (1+r_0)\kappa\, \lambda(\widetilde{u}_n)$, we see that

$$\|\mathbf{P}^n u - \lambda(u) h\|_\infty \leq \frac{\sup h + (1+r_0)\kappa}{(1-p)}\; \lambda(u)\, \lambda(1-p)^n \qquad \text{for } n \geq 1,$$

implying the assertion of the theorem. To establish our claim, take $\kappa := \kappa(r_0)$ as in Lemma 3 and choose $p \in (0,1)$ so small that

$$p(1+r_0)^2\kappa^2 < 1 \quad \text{ and } \quad p\, \frac{2r_0(1+r_0)^2\kappa^2}{1-p(1+r_0)^2\kappa^2} \leq (1-\rho_0)(r_0 - A_0).$$

Assume then that (4) holds for $k = n-1$. By (the second estimate of) Lemma 1, $\widetilde{u}_n$ is β-admissible with $R_\beta(\widetilde{u}_n) \leq 2A_0$. According to Lemma 3, $\kappa^{-1} \leq \lambda_B(\widetilde{u}_n)/\lambda(\widetilde{u}_n) \leq \kappa$ for $B \in \beta$. Since $u_n = \triangle_{p,h}(\widetilde{u}_n)$ and $p(1+R_\beta(\widetilde{u}_n))(1+R_\beta(h))\kappa^2 \leq p(1+r_0)^2\kappa^2 < 1$, Lemma 2 implies that u_n is β-admissible with

$$R_\beta(u_n) \leq R_\beta(\widetilde{u}_n) + p\frac{2r_0(1+r_0)^2\kappa^2}{1-p(1+r_0)^2\kappa^2}.$$

Therefore, by the first estimate in Lemma 1 and our choice of p,

$$R_\beta(u_n) \leq \rho_0 r_0 + A + (1-\rho_0)(r_0 - A) = r_0,$$

as required.

Finally, if $R_\beta(u) > r_0$, recall that Lemma 1 implies $R_\beta(\mathbf{P}^n u) \leq \rho_0^n R_\beta(u) + A_0$ for $n \geq 1$. This provides us with $n_1 \geq 1$ so large that $u_0 := \mathbf{P}^{n_1} u$ satisfies $R_\beta(u_0) \leq r_0$, and the previous argument applies. ■

Remark 4 (Probabilistic coupling) *Up to this point we have been using the term* coupling *synonymously for our procedure of removing overlapping portions of mass. To see how this is related to the probabilistic concept of coupling random variables, recall that the latter simply means to choose a joint distribution of the variables, compare [BL] or [Lin]. The random variables to be coupled here are the random elements* $\mathrm{X}_n := T^n \mathrm{X}_0$ *and* $\mathrm{Y}_n := T^n \mathrm{Y}_0$, $n \geq 0$, *of the space* X, *where* X_0 *and* Y_0 *are distributed according to the probability densities* h *and* u *on* $(X, \mathcal{B}, \lambda)$. *The decomposition of the density of* X_n *obtained in our proof,* $\mathbf{P}^n u = u_n + (1-(1-p)^n)\, h$, *immediately gives rise to a coupling of* X_n *and* Y_n *on the product space: Take the common part of* $\mathbf{P}^n u$ *and* $\mathbf{P}^n h = h$ *and place it on the diagonal* $D \subseteq X^2$ *to obtain a singular measure* $P'_n := \pi_D^{-1}((1-(1-p)^n)\, h \cdot d\lambda)$, *where* π_D *denotes the projection of* D *onto the first coordinate. For the remaining mass, choose any coupling* P''_n *you like, e.g. the independent one with density* $(x,y) \longmapsto (1-p)^{2n} h(x)\, u_n(y)$ *w.r.t.* $\lambda \times \lambda$. *Then* $P_n := P'_n + P''_n$ *is a coupling of* X_n *and* Y_n *whose dominant correlated part* P'_n *reflects the close approach of the respective distributions.*

Remark 5 (Coupling pairs of densities) *We could of course start with any pair of admissible initial densities, couple them to each other (as below) to see that their difference decreases exponentially fast, and derive the existence of an admissible invariant density as a corollary of this.*

5 Reformulation in terms of cones

We review the core of the above argument in more abstract terms, showing that it is equivalent to the cone method (cf. [Ba], [Liv]). Let $\mathcal{C}$ be a proper convex cone in the real linear space V, i.e. $\mathcal{C} \subseteq V \setminus \{0\}$ with $p\mathcal{C} \subseteq \mathcal{C}$ for any $p > 0$, and $u + v \in \mathcal{C}$

whenever $u, v \in \mathcal{C}$. Assume that $R : \mathcal{C} \to (0,\infty)$ satisfies $R(pu) = R(u)$ for any $p > 0$ and $u \in \mathcal{C}$, as well as $R(u+v) \leq \max(R(u), R(v))$ for $u, v \in \mathcal{C}$. Then, for any $r > 0$, $\mathcal{C}_r := \{u \in \mathcal{C} : R(u) \leq r\}$ defines a convex cone in $\mathcal{C}$. In the framework of the preceding section, we took $R := R_\beta$ and $\mathcal{C} := \{u$ β-admissible on $X : R_\beta(u) \leq r_0\}$.

Suppose that $\mathcal{L} : \mathcal{C} \to \mathcal{C}$ is a linear operator (corresponding to the transfer operator $\mathbf{P}$) satisfying a Doeblin-Fortet type inequality with respect to R, i.e.

$$R(\mathcal{L}u) \leq \rho R(u) + A \qquad \text{for } u \in \mathcal{C}, \tag{5}$$

with $\rho \in (0,1)$ and $A > 0$. Then $\mathcal{L}(\mathcal{C}_r) \subseteq \mathcal{C}_r$ for $r \geq A_1 := (1-\rho)^{-1}A$. Moreover, for any $\gamma \in (\rho, 1)$ we have $\mathcal{L}(\mathcal{C}_r) \subseteq \mathcal{C}_{\gamma r}$ provided that $r \geq A_\gamma := [\gamma(\gamma - \rho)]^{-1}A$.

Let $\lambda : \mathcal{C} \to (0,\infty)$ be a linear functional, then $\mathcal{B} := \mathcal{C} \cap \{\lambda = 1\}$ is a base of the cone $\mathcal{C}$, i.e. every ray $\{p\, u : p > 0\}$, $u \in \mathcal{C}$, intersects B in a unique point $\widetilde{u}$. As before we define the coupling operator $\triangle_{p,v}(u) := u - p\,(\lambda(u)/\lambda(v))v$, $p > 0$, $u, v \in \mathcal{C}$. The λ-proportion of u which can be coupled with v preserving some R-bound $r > 0$ is given by $\vartheta_r(u,v) := \sup\{p > 0 : \triangle_{p,v}(u) \in \mathcal{C}_r\} = \sup\{p > 0 : \widetilde{u} - p\widetilde{v} \in \mathcal{C}_r\}$.

Fix $\gamma \in (\rho, 1)$ and $r \geq A_\gamma$, and assume that

$$p := \inf_{u,v \in \mathcal{L}(\mathcal{C}_r)} \vartheta_r(u,v) > 0. \tag{6}$$

Then obviously $\triangle_{p,v}(\mathcal{L}(\mathcal{C}_r)) \subseteq \mathcal{C}_r$, so that for any pair $(u,v) = (u_0, v_0) \in \mathcal{C}_r \times \mathcal{C}_r$ with $\lambda(u) = \lambda(v)$ the inductive scheme $u_{n+1} := \triangle_{p,\mathcal{L}(v_n)}(\mathcal{L}(u_n))$ and $v_{n+1} := (1-p)\mathcal{L}(v_n)$, $n \geq 1$, defines sequences in $\mathcal{C}_r$ with $\mathcal{L}^n(u - v) = u_n - v_n$ and $\lambda(u_n) \leq (1-p)^n \lambda(u)$, $n \geq 1$. This is the coupling scheme used above (with the special choice $v := h$). In the situation of the previous section, (6) follows from Lemmas 3 and 2, which ensure that $\mathbf{P}(\mathcal{C}_r) \subseteq \{u \in \mathcal{C} : \kappa(r_0)^{-1} \leq \lambda_B(u)/\lambda(u) \leq \kappa(r_0)$ for all $B \in \beta\}$, and that for any $r > 0$, $u, v \in \mathbf{P}(\mathcal{C}_r)$, and $p > 0$ sufficiently small, $R(\widetilde{u} - p\widetilde{v}) \leq R(u) + d_r(p)$, with $\lim_{p \to 0} d_r(p) = 0$.

The crucial condition (6) is equivalent to requiring that $\mathcal{L}(\mathcal{C}_r) \subseteq \mathcal{C}_{\gamma r}$ should be a bounded subset of $\mathcal{C}_r$ equipped with the Hilbert pseudo-metric $\Theta_r(u,v) := -\log(\vartheta_r(u,v)/\vartheta_r(v,u))$, $u, v \in \mathcal{C}_r$, and its consequence $\triangle_{p,v}(\mathcal{L}(\mathcal{C}_r)) \subseteq \mathcal{C}_r$ which makes the coupling procedure possible means that $\mathcal{L}$ strictly contracts Θ_r, which is usually seen via Birkhoff's inequality (cf. [Ba] or [Liv]).

Acknowledgments. I am indebted to M.Thaler for comments and suggestions which significantly improved the presentation. This research was supported by the Austrian Science Foundation FWF, project P14734-MAT.

References

[Aa] J.Aaronson: *An introduction to infinite ergodic theory.* American Mathematical Society, Mathematical Surveys and Monographs, Vol. 50, 1997.

[Ba] V.Baladi: *Positive transfer operators and decay of correlations.* World Scientific 2000.

[Be] A.Berechet: *A Kuzmin-type theorem with exponential convergence for a class of fibred systems.* Ergod. Th. & Dynam. Sys. 21 (2001), 673-688.

[BFG] X.Bressaud, R.Fernandez, A.Galves: *Decay of correlation for non-Hölderian dynamics. A coupling approach.* Electron. J. Prob. 4 (1999), 19pp.

[BL] X.Bressaud, C.Liverani: *Anosov diffeomorphisms and coupling.* Ergod. Th. & Dynam. Sys. 22 (2002), 129-152.

[Br] A.Broise: *Fractions continues multidimensionelles et lois stables.* Bull. Soc. Math. France 124 (1996), 97-139.

[Ga] C.F.Gauss: *Letter to Laplace.* Göttingen, January 30*th* 1812.

[Ho] F.Hofbauer: *Piecewise invertible dynamical systems.* Probab. Th. Rel. Fields 72 (1986), 359-386.

[Ke] G.Keller: *Lifting measures to markov extensions.* Mh. Math. 108 (1989), 183-200.

[Ku] R.O.Kuzmin: *Sur un problème de Gauss.* Atti Congr. Int. Bologne 6 (1928), 83-89.

[Le] P.Lévy: *Sur le loi de probabilité dont dependents les quotients complets et incomplets d'une fraction continue.* Bull. Soc. Math. France 557 (1929), 178-194.

[Lin] T.Lindvall: *Lectures on the coupling method.* Wiley 1992.

[Liv] C.Liverani: *Decay of correlations.* Ann. Math. (2) 142 (1995), 239-301.

[Me] D.H.Meyer: *Approach to equilibrium for locally expanding maps in* $\mathbb{R}^k$. Commun. Math. Phys. 95 (1984), 1-15.

[PY] G.Pianigiani, J.A.Yorke: *Expanding maps on sets which are almost invariant: Decay and chaos.* Trans. Amer. Math. Soc. 252 (1979), 351-366.

[S1] F.Schweiger: *Ergodic theory of fibered systems and metric number theory.* Clarendon Press, Oxford 1995.

[S2] F.Schweiger: *Kuzmin's theorem revisited.* Ergod. Th. & Dynam. Sys. 20 (2000), 557-565.

[S3] F.Schweiger: *Multidimensional continued fractions.* Oxford UP, 2000.

[Y1] L.S.Young: *Statistical properties of dynamical systems with some hyperbolicity.* Ann. Math. 147 (1998), 585-650.

[Y2] L.S.Young: *Recurrence times and rates of mixing.* Isr. J. Math. 110 (1999), 153-188.

[Yu] M.Yuri: *Multi-dimensional maps with infinite invariant measures and countable state sofic shifts.* Indag. Math., N.S. 6 (3) (1995), 355-383.

[Zw] R.Zweimüller: *Ergodic properties of infinite measure preserving interval maps with indifferent fixed points.* Ergod. Th. & Dynam. Sys. 20 (2000), 1519-1549.