Ergod. Th. & Dynam. Sys. (2011), **31**, 259–284 doi:10.1017/S0143385709001060
© *Cambridge University Press, 2010*

Polymorphisms and adiabatic chaos

A. NEISHTADT†‡ and D. TRESCHEV§

† *Department of Mathematical Sciences, Loughborough University, Loughborough, LE11 3TU, UK*

‡ *Space Research Institute, Profsoyuznaya 84/32, Moscow, 117997, Russia*

§ *Steklov Mathematical Institute, Gubkina 8, Moscow, 119991, Russia*

(*Received* 14 *August* 2008 *and accepted in revised form* 28 *August* 2009)

Abstract. At the end of the last century Vershik introduced some dynamical systems, called polymorphisms. Systems of this kind are multivalued self-maps of an interval, where (roughly speaking) each branch has some probability. By definition, the standard Lebesgue measure should be invariant. Unexpectedly, some class of polymorphisms appeared in the problem of destruction of an adiabatic invariant after a multiple passage through a separatrix. We discuss ergodic properties of polymorphisms from this class.

1. *Introduction*

In [**7**] and [**8**] Vershik introduced some dynamical systems called polymorphisms. Any polymorphism is a dynamical system determined by a multivalued self-map of a measure space (M, μ). We will consider the case† when M is the interval $[0, 1]$ and μ is the standard Lebesgue measure dx.

A polymorphism T is defined as the ordered diagram

$$([0, 1]_x, dx) \xleftarrow{\pi_x} ([0, 1]_x \times [0, 1]_y, \nu) \xrightarrow{\pi_y} ([0, 1]_y, dy), \tag{1.1}$$

where π_x and π_y are projections to the x-component and y-components respectively of the product $[0, 1]_x \times [0, 1]_y$ and ν is a probability measure such that

$$\pi_x \nu = dx \quad \text{and} \quad \pi_y \nu = dy. \tag{1.2}$$

The dynamical interpretation of T is as follows: T maps randomly any measurable set $\Lambda \subset [0, 1]_x$ to the interval $[0, 1]_y$ so that the probability of a measurable set $\Omega \subset [0, 1]_y$ equals $\nu(\Lambda \times \Omega)$.

Unexpectedly, polymorphisms appeared in the problem of destruction of an adiabatic invariant after a multiple passage through a separatrix. However the measures ν, appearing

† Recall that, if (M, μ) is a Lebesgue space, where μ is a measure without atoms, it is metrically isomorphic to $([0, 1], dx)$.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

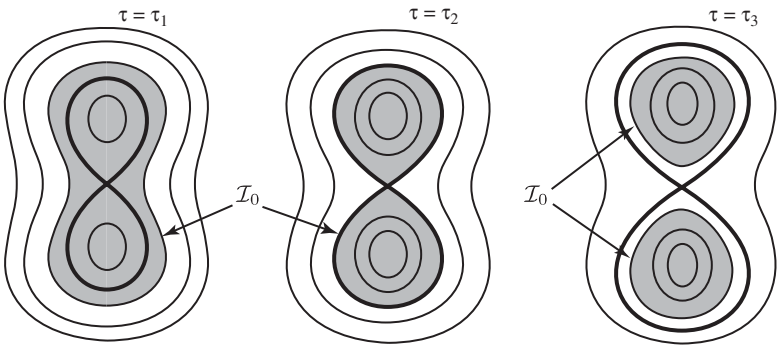


FIGURE 1. Phase portraits of the 'frozen' system. Grey domains correspond to the same value $\mathcal{I}_0$ of the adiabatic invariant.

in this way, have a special structure: they are supported on several curves, lying in $[0, 1]_x \times [0, 1]_y$. Below we define this class of polymorphisms and discuss ergodic properties of some its representatives.

Our plan is as follows. In §2 we explain how polymorphisms appear in the problem of adiabatic chaos. This section is independent of the remaining part of the paper and a reader, not interested in motivations, in principle can skip it. In §3 we define the class of polymorphisms which is the subject of this paper and present basic constructions including the Perron–Frobenius operator, the concepts of ergodicity and mixing. In §4 we define some important subclass of polymorphisms and present several existence statements for this class. In §5 we present examples of polymorphisms with non-ergodic dynamics. In §6 we discuss metric properties of the Perron–Frobenius operator for polymorphisms from §4. As a corollary we obtain two theorems which present sufficient conditions for mixing. In §7 we study ergodic properties of a two-parametric family of piecewise-linear polymorphisms $T_{\beta,s}$. Appendices contain proofs of some technical statements.

2. *Motivation*

Consider the Hamiltonian system with one and a half degrees of freedom

$$\dot{x} = \partial H/\partial y, \quad \dot{y} = -\partial H/\partial x, \quad H = H(x, y, \varepsilon t),$$

where dots are derivatives with respect to time t, H is $\hat{\tau}$-periodic in $\tau = \varepsilon t$,

$$H(x, y, \tau) = H(x, y, \tau + \hat{\tau}),$$

and ε is a small parameter.

First, consider the case of fixed ('frozen') τ. Then we have a system with one degree of freedom (the 'frozen system'). Suppose that for any fixed τ the level lines of H (the phase portrait of the frozen system) are as in Figure 1. More precisely, we assume that for any τ there is a hyperbolic fixed point $\hat{z}(\tau)$ with two homoclinic separatrix loops ('upper' $\gamma_+(\tau)$ and 'lower' $\gamma_-(\tau)$), and the picture depends smoothly on τ.

For small $\varepsilon \neq 0$ the system has an adiabatic invariant: the action $\mathcal{I}$ in the 'frozen system'. The separatrices divide the phase space $\{x, y\}$ of the 'frozen system' into three

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

domains: $D_+(\tau)$, $D_-(\tau)$, and $D_0(\tau)$ with the boundaries

$$\partial D_+ = \gamma_+ \cup \hat{z}, \quad \partial D_- = \gamma_- \cup \hat{z}, \quad \partial D_0 = \gamma_+ \cup \gamma_- \cup \hat{z}.$$

In each of these domains the action $\mathcal{I} = \mathcal{I}(H, \tau)$ equals the area inside the curve $H =$ constant, see [**1**]. It is smooth outside the separatrix and discontinuous on the separatrix. For a motion far from the separatrices the value of $\mathcal{I}$ for a moving phase point remains constant with accuracy of order ε on time intervals of order $1/\varepsilon$. Approximation in which the value of $\mathcal{I}$ is considered as a constant in the process of the motion is called *the adiabatic approximation*.

For the case when a phase point crosses the separatrices the adiabatic approximation should be modified as follows. The value of $\mathcal{I}$ is considered as a constant up to an arrival of the phase point at a separatrix. At this time moment the phase point can change the domain. Different phase points which arrive at the separatrix simultaneously can be captured into different domains. A certain probability can be assigned to the capture into each domain (see below). After the capture into a certain domain the value of the action is considered as a constant up to the next arrival at the separatrix. Estimates of the accuracy of this approximation can be found in [**4**]. This form of the adiabatic approximation allows us to describe the dynamics as an evolution of a set of 'adiabatic' phase points. Each of the phase points in this set has a value of $\mathcal{I}$ calculated by means of the adiabatic approximation and a certain probability assigned to it.

Now we define a multivalued map $\mathcal{I} \mapsto T(\mathcal{I})$ which associates with $\mathcal{I}$ at the time moment $\tau = 0$ its value $T(\mathcal{I})$ at the time moment $\tau = \hat{\tau}$ in the adiabatic approximation.

First, we determine the domain, where the map T is defined. Let

$$A_+(\tau) = \int_{\gamma_+(\tau)} y\,dx, \quad A_-(\tau) = \int_{\gamma_-(\tau)} y\,dx,$$
$$A_0(\tau) = A_+(\tau) + A_-(\tau) \tag{2.1}$$

be the areas bounded by the upper, lower separatrix loops and of all the separatrix respectively. These quantities can be regarded as areas of the domains D_+, D_-, and the complement of D_0 respectively.

Let $U(\tau)$ be the disjoint union of three intervals

$$U_\pm(\tau) = (0,\, A_\pm(\tau)), \quad U_0(\tau) = (A_0(\tau),\, C),$$

where $C = \max_\tau A_0(\tau)$. We consider $U(\tau)$ together with the standard Lebesgue measure du. By (2.1) the total measure

$$|U(\tau)| = A_+(\tau) + A_-(\tau) + C - A_0(\tau) = C$$

does not depend on τ. Here we assume that if $\mathcal{I} \in U_+(\tau)$ ($\mathcal{I} \in U_-(\tau)$, $\mathcal{I} \in U_0(\tau)$), the corresponding phase point lies in $D_+(\tau)$ ($D_-(\tau)$, $D_0(\tau)$).

The domain of T is $U(0)$. For an arbitrary partition

$$0 = \tau_0 < \tau_1 < \cdots < \tau_N = \hat{\tau} \tag{2.2}$$

the map T is the composition

$$T = T_{\tau_{N-1}}^{\tau_N} \circ \cdots \circ T_{\tau_1}^{\tau_2} \circ T_{\tau_0}^{\tau_1}, \tag{2.3}$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

where $T_{\tau_s}^{\tau_{s+1}} : U(\tau_s) \to U(\tau_{s+1})$ associates with $\mathcal{I} = \mathcal{I}|_{\tau=\tau_s}$ the quantity $\mathcal{I}|_{\tau=\tau_{s+1}} = T_{\tau_s}^{\tau_{s+1}}(\mathcal{I})$. To define this map, we will have to consider several cases.

(a) Suppose that $\mathcal{I} \in U_*(\tau_s)$, $* \in \{+, -, 0\}$, and the equation

$$A_*(\tau) = \mathcal{I}, \quad \tau \in [\tau_s, \tau_{s+1}]$$

has no solution. Then $T_{\tau_s}^{\tau_{s+1}}(\mathcal{I}) = \mathcal{I}$. This corresponds to the fact that $\mathcal{I}$ is preserved if the phase point does not cross the separatrix.

In other cases the phase points cross the separatrix and we will assume that the interval $[\tau_s, \tau_{s+1}]$ is so small that this event takes place only once on $[\tau_s, \tau_{s+1}]$. Moreover, for convenience we assume that this happens at some time moment $\tilde{\tau} \in (\tau_s, \tau_{s+1})$.

(b_0) Suppose that $\mathcal{I} \in U_0(\tau_s)$ and

$$A_0(\tilde{\tau}) = \mathcal{I} \quad \text{for some } \tilde{\tau} \in (\tau_s, \tau_{s+1}).$$

This means that the area $A_0(\tau)$ increases and for $\tau = \tilde{\tau}$ becomes equal to $\mathcal{I}$. In particular, $A_0'(\tilde{\tau}) \geq 0$. (Here and below 'prime' denotes derivative with respect to τ.) For $\tau > \tilde{\tau}$ the phase points fall into the domains D_+ and D_-:

$$T_{\tau_s}^{\tau_{s+1}}(\mathcal{I}) = \begin{cases} A_+(\tilde{\tau}) \in U_+(\tau_{s+1}) & \text{with probability } p_0^+(\tilde{\tau}), \\ A_-(\tilde{\tau}) \in U_-(\tau_{s+1}) & \text{with probability } p_0^-(\tilde{\tau}), \end{cases}$$

where

$$p_0^+(\tilde{\tau}) = \frac{A_+'(\tilde{\tau})}{A_0'(\tilde{\tau})}, \quad p_0^-(\tilde{\tau}) = \frac{A_-'(\tilde{\tau})}{A_0'(\tilde{\tau})} \text{ provided } A_\pm'(\tilde{\tau}) \geq 0.$$

If $A_+'(\tilde{\tau})$ or $A_-'(\tilde{\tau})$ is negative, the phase points fall into the domain bounded by the increasing loop with probability 1.

(b_+) Suppose that $\mathcal{I} \in U_+(\tau_s)$ and

$$A_+(\tilde{\tau}) = \mathcal{I} \quad \text{for some } \tilde{\tau} \in (\tau_s, \tau_{s+1}).$$

This means that the area $A_+(\tau)$ decreases and for $\tau = \tilde{\tau}$ becomes equal to $\mathcal{I}$. In particular, $A_+'(\tilde{\tau}) \leq 0$. The phase points leave D_+ and for $\tau > \tilde{\tau}$ are captured into D_0 and D_-:

$$T_{\tau_s}^{\tau_{s+1}}(\mathcal{I}) = \begin{cases} A_0(\tilde{\tau}) \in U_0(\tau_{s+1}) & \text{with probability } p_+^0(\tilde{\tau}), \\ A_-(\tilde{\tau}) \in U_-(\tau_{s+1}) & \text{with probability } p_+^-(\tilde{\tau}), \end{cases}$$

where

$$p_+^0(\tilde{\tau}) = \frac{-A_0'(\tilde{\tau})}{-A_+'(\tilde{\tau})}, \quad p_+^-(\tilde{\tau}) = \frac{A_-'(\tilde{\tau})}{-A_+'(\tilde{\tau})} \text{ provided } A_0'(\tilde{\tau}) < 0 \text{ and } A_-'(\tilde{\tau}) > 0.$$

If $A_0'(\tilde{\tau}) > 0$, the phase points fall into D_-, and if $A_-'(\tilde{\tau}) < 0$, into D_0 with probability 1.

(b_-) The case $\mathcal{I} \in U_-(\tau_s)$ is analogous to (b_+).

Obviously, T, defined by (2.3), does not depend on partition (2.2).

Dynamical systems, generated by these multivalued maps are expected to have generically strong ergodic properties which implies a fast stochastization in the original Hamiltonian system if $\mathcal{I}_0 \in U(0)$.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

Remark 2.1. If $A_+(\tau) \equiv A_-(\tau)$, the map T is, in a sense, trivial:

$$T(\mathcal{I}) = \begin{cases} \mathcal{I} \in U_0(0) & \text{if } \mathcal{I} \in U_0(0), \\ \mathcal{I} \in U_+(0) \text{ with probability } 1/2 & \text{if } \mathcal{I} \in U_\pm(0), \\ \mathcal{I} \in U_-(0) \text{ with probability } 1/2 & \text{if } \mathcal{I} \in U_\pm(0). \end{cases}$$

Chaos in this situation is produced by a more delicate effect: small error terms appeared because of scattering on the separatrix (recall that these terms are skipped in the adiabatic approximation we use here) [**2, 5, 6**].

The multi-valued map T can be regarded as a collection of maps $\varphi_j : I_j \to U(0)$, $j = 1, \dots, J$, where $I_j \subset U(0)$ are intervals. Any point $u \in U(0)$ lying in $I_{s_1}, \dots, I_{s_k}$ has the images $\varphi_{s_1}(u), \dots, \varphi_{s_k}(u)$ with probabilities $p_{s_1}(u), \dots, p_{s_k}(u)$ respectively. Without loss of generality we can assume that φ_j are monotone because in case of necessity we can divide I_j by subintervals of monotonicity of the φ_j.

THEOREM 1. *The map T preserves the measure du on $U(0)$: for almost all $u \in U(0)$*

$$\sum_{l=1}^{k} \frac{p_{s_l}(v_l)}{|\varphi'_{s_l}(v_l)|} = 1, \tag{2.4}$$

where $v_1, \dots, v_k$ are all preimages of u: $\varphi_{s_l}(v_l) = u$.

COROLLARY 2.1. *For any measurable set $S \in U(0)$ let $S_1, \dots, S_J$ be its preimage: $S_j = \{w \in I_j \mid \varphi_j(w) \in S\}$. Then*

$$\int_S du = \sum_{j=1}^{J} \int_{S_j} p_j(v)\, dv.$$

We prove Theorem 1 in Appendix A.

According to the terminology, introduced by Vershik, preservation of the Lebesgue measure means that T is a *polymorphism*. Having in mind this motivation, below we study polymorphisms from dynamical point of view.

Example. Here is an example of a polymorphism associated with separatrix crossings. To have simple formulas we consider the case when the areas $A_\pm(\tau)$ of the domains $D_\pm(\tau)$ are piecewise linear functions of τ.

In what follows A_*, $\hat\tau$, τ_1, d, γ_0, γ_1, γ_3 are positive constants, and γ_2 is a non-negative constant, $\tau_1 + d < \hat\tau$, $\gamma_2 < \gamma_1$, $\gamma_3 < \gamma_1$. Assume that at $\tau = 0$ both areas $A_\pm$ have the same value A_*. Assume that for $0 \le \tau \le \tau_1$ both functions $A_\pm(\tau)$ grow linearly with the same rate γ_0: $A'_\pm(\tau) = \gamma_0$. For $\tau_1 \le \tau \le \hat\tau - d$ the function $A_+(\tau)$ decays linearly with the rate γ_1: $A'_+(\tau) = -\gamma_1$, and for $\hat\tau - d \le \tau \le \hat\tau$ it decays linearly with the rate γ_2: $A'_+(\tau) = -\gamma_2$. For $\tau_1 \le \tau \le \hat\tau$ the function $A_-(\tau)$ decays linearly with the rate γ_3: $A'_-(\tau) = -\gamma_3$. The constants are such that $A_\pm(\hat\tau) = A_\pm(0) = A_*$, i.e.,

$$\gamma_0\tau_1 - \gamma_1(\hat\tau - d - \tau_1) - \gamma_2 d = 0 \quad \text{and} \quad \gamma_0\tau_1 - \gamma_3(\hat\tau - \tau_1) = 0.$$

Consider the motion of phase points with some initial, at $\tau = 0$, action $\mathcal{I} \in [A_*, C]$, $C = A_* + 2\gamma_0\tau_1$. At $\tau = 0$ these phase points are placed in the domain $D_0(0)$. During the

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

motion they cross the separatrix at some moment of time and are captured into D_+ and D_-. Then they cross the separatrix again and return into D_0 with the values of the action $\mathcal{I}_+$ (for phase points captured into D_+) and $\mathcal{I}_-$ (for phase points captured into D_-). The probability of capture into each of the domains $D_\pm(\tau)$ is $1/2$. Description of the motion in the adiabatic approximation leads to the following formulas (we omit straightforward calculations):

$$\mathcal{I}_+ = \begin{cases} A_* + \dfrac{(\mathcal{I} - A_*)(\gamma_3 + \gamma_2)}{2\gamma_2} & \text{if } \mathcal{I} \in [A_*,\, A_* + 2\gamma_2 d], \\ A_* + \dfrac{(\mathcal{I} - A_*)(\gamma_3 + \gamma_1)}{2\gamma_1} + \dfrac{\gamma_3(\gamma_1 - \gamma_2)d}{\gamma_1} & \text{if } \mathcal{I} \in [A_* + 2\gamma_2 d,\, C], \end{cases}$$

and

$$\mathcal{I}_- = \begin{cases} A_* + \dfrac{(\mathcal{I} - A_*)(\gamma_3 + \gamma_2)}{2\gamma_3} & \text{if } \mathcal{I} \in [A_*,\, A_* + 2\gamma_3 d], \\ A_* + \dfrac{(\mathcal{I} - A_*)(\gamma_3 + \gamma_1)}{2\gamma_3} + (\gamma_2 - \gamma_1)d & \text{if } \mathcal{I} \in [A_* + 2\gamma_3 d,\, C]. \end{cases}$$

These formulas define a polymorphism T. The domain of T is $[A_*, C]$; T maps $\mathcal{I}$ to $\mathcal{I}_+$ and $\mathcal{I}_-$ with probabilities $1/2$.

Introducing the new notation

$$x = \frac{\mathcal{I} - A_*}{C - A_*}, \quad s = \frac{d(\gamma_2 + \gamma_3)}{C - A_*}, \quad \alpha = \frac{2\gamma_2}{\gamma_2 + \gamma_3}, \quad \beta = \frac{2\gamma_3}{\gamma_2 + \gamma_3}$$

we arrive at the following equations, determining T:

$x \in [0, 1]$ is transformed to $\varphi_1(x)$ or to $\varphi_2(x)$ with probabilities $1/2$,

$$\varphi_1 = \begin{cases} x/\beta & \text{if } 0 \le x \le \beta s, \\ \dfrac{1 - s}{1 - \beta s} x + \dfrac{s - \beta s}{1 - \beta s} & \text{if } \beta s \le x \le 1, \end{cases} \tag{2.5}$$

$$\varphi_2 = \begin{cases} x/\alpha & \text{if } 0 \le x \le \alpha s, \\ \dfrac{1 - s}{1 - \alpha s} x + \dfrac{s - \alpha s}{1 - \alpha s} & \text{if } \alpha s \le x \le 1. \end{cases} \tag{2.6}$$

The coefficients α and β enter these formulas symmetrically and $\alpha + \beta = 2$. Hence we can assume that $0 < \beta \le 1$. We have a two-parametric family of polymorphisms

$$T_{\beta,s}, \quad 0 < \beta \le 1,\ 0 < s < 1,\ (2 - \beta)s < 1.$$

3. *Basic definitions*

In this section we define a specific class of polymorphisms of an interval with the Lebesgue measure. Below without loss of generality we assume that this interval is $[0, 1]$. According to argument from §2 random multi-valued maps, responsible for jumps of the adiabatic invariant at the moments of the separatrix crossing are polymorphisms from this class. Here an interesting question appears. *Which polymorphisms, defined in §3, can be generated by Hamiltonian systems from §2?*

We plan to discuss this question elsewhere. The majority of definitions and statements from this section have natural analogs in the general situation, [**7, 8**].

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

3.1. *Multivalued self-maps of an interval.* Let the interval $[0, 1]$ be presented as the union

$$[0, 1] = \bigcup_{j=1}^{J} I_j, \quad I_j = [a_j, b_j].$$

The intervals I_j can have non-trivial pairwise intersections. For example, it can happen that all I_j equal $[0, 1]$. Below, j always takes values from the set $\{1, \dots, J\}$. For any j consider the functions

$$\varphi_j : I_j \to [0, 1], \quad p_j : I_j \to [0, 1].$$

We assume that for any j the functions φ_j and p_j are smooth or piecewise smooth inside I_j.

For any $x \in [0, 1]$ we put

$$V(x) = \{j \mid x \in I_j\}.$$

We assume that the following conditions hold.

P (Probability) $\sum_{j \in V(x)} p_j(x) = 1$ for any $x \in [0, 1]$.

M (Monotonicity) the functions φ_j are strictly monotone on I_j.

According to **M** there exist the inverse functions $\psi_j = \varphi_j^{-1}$ which are smooth or piecewise smooth on $K_j = \varphi_j(I_j)$.

Consider the following dynamical system T on $[0, 1]$. Any point $x \in I_j$ is mapped to $\varphi_j(x)$ with probability $p_j(x)$. We write

$$T = (\varphi_1, \dots, \varphi_J; p_1, \dots, p_J; I_1, \dots, I_J)$$

or (in shorter notation) $T = (\varphi; p; I)$.

3.2. *The operator W_T.* Consider the space $L_2 = L_2([0, 1], dx)$ of functions, square integrable on $[0, 1]$ with respect to the Lebesgue measure dx. Below we use the notation $\langle , \rangle$ for the corresponding scalar product and $\| \cdot \|$ for the corresponding norm. In a standard way we define the Perron–Frobenius operator

$$W_T : L_2 \to L_2, \quad f \mapsto W_T f.$$

For any $y \in [0, 1]$ we put

$$V^*(y) = \{j \mid y \in \varphi_j(I_j)\}.$$

Then $\{\psi_j(y) \mid j \in V^*(y)\}$ is the set of all preimages of the point y with respect to T. By definition

$$W_T f(y) = \sum_{j \in V^*(y)} p_j \circ \psi_j(y) |\psi_j'(y)| f \circ \psi_j(y). \tag{3.1}$$

For any measurable set $\Omega \subset [0, 1]$ we have

$$\int_\Omega W_T f(y)\, dy = \sum_{j=1}^{J} \int_{\varphi_j^{-1}(\Omega)} p_j(x) f(x)\, dx.$$

The positive cone

$$L_2^+ = \{\rho \in L_2 \mid \rho \geq 0\}$$

can be associated with the space of densities of finite measures on $[0, 1]$. We have the obvious inclusion

$$W_T(L_2^+) \subset L_2^+. \tag{3.2}$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

3.3. *Polymorphisms.* Below we assume that the Lebesgue measure is invariant.

L (Lebesgue) $W_T 1 = 1$.

Equation (3.1) implies that

$$\sum_{j \in V^*(y)} p_j \circ \psi_j(y) |\psi_j'(y)| = 1. \tag{3.3}$$

COROLLARY 3.1. *For any* $y \in [0, 1] V^*(y) \neq \emptyset$.

Definition 3.1. Any map $(\varphi; p; I)$, satisfying **P**, **M**, and **L**, will be said to be a *polymorphism*.

A polymorphism can be regarded as a multivalued self-map of an interval preserving the Lebesgue measure.

The following construction shows a connection between Vershik's definition of a polymorphism (see the introduction) and Definition 3.1. Suppose that $(\varphi; p; I)$ is a polymorphism in the sense of Definition 3.1. Let ν be the following measure, supported on the graphs of the functions φ_j. For any $S \subset [0, 1] \times [0, 1]$ let $\chi_S : [0, 1] \times [0, 1] \to \mathbb{R}$ be its indicator:

$$\chi_S(x, y) = \begin{cases} 0 & \text{if } (x, y) \notin S, \\ 1 & \text{if } (x, y) \in S. \end{cases}$$

Then by definition

$$\nu(S) = \sum_{j=1}^{J} \int_{I_j} p_j(x) \chi_S(x, \varphi_j(x))\, dx.$$

We have an obvious proposition.

PROPOSITION 3.1. *The measure* ν *satisfies (1.2).*

3.4. *An adjoint polymorphism.* For any polymorphism $T = (\varphi; p; I)$ we put

$$K_j = \varphi_j(I_j), \quad q_j(y) = p_j \circ \psi_j(y) |\psi_j'(y)|.$$

PROPOSITION 3.2. $(\psi; q; K)$ *is a polymorphism.*

Proof. The equation $[0, 1] = \bigcup_{j=1}^{J} K_j$ follows from Corollary 3.1. Condition **M** holds because φ_j are smooth. We have

$$\sum_{j=1}^{J} q_j(y) = \sum_{j=1}^{J} p_j \circ \psi_j(y) |\psi_j'(y)| = 1,$$

$$\sum_{j=1}^{J} q_j \circ \varphi_j(x) |\varphi_j'(x)| = \sum_{j=1}^{J} p_j(x) = 1.$$

Hence Conditions **P** and **L** also hold. □

In fact it was sufficient to note that the passage from $(\varphi; p; I)$ to $(\psi; q; K)$ corresponds to the exchange of left and right in diagram (1.1).

We say that $(\psi; q; K)$ is adjoint to T: $(\psi; q; K) = T^*$. Obviously $T^{**} = T$.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

PROPOSITION 3.3. *The operator W_{T^*} is adjoint to W_T with respect to the L_2-scalar product: $W_{T^*} = W_T^*$.*

Proof. Straightforward computation. □

COROLLARY 3.2. *For any polymorphism T we have $W_T^* 1 = 1$.*

3.5. *Structures.* Condition **L** combined with (3.2) and Corollary 3.2 mean that W_T is a Markov (bistochastic) operator, i.e., W_T satisfies the following conditions.

(1) $W_T(L_2^+) \subset L_2^+$.

(2) $W_T 1 = W_T^* 1 = 1$.

PROPOSITION 3.4. *The operator W_T is a contraction on L_2, i.e.,*

$$\|W_T f\| \le \|f\| \quad \textit{for any } f \in L_2([0, 1]).$$

Proof. Due to (1) and the equation $W_T 1 = 1$ we can use the Jensen inequality:

$$(W_T f(x))^2 \le W_T f^2(x).$$

Therefore

$$\begin{aligned} \|W_T f\|^2 &= \int_0^1 (W_T f(x))^2 \, dx \le \int_0^1 W_T f^2(x) \, dx \\ &= \langle W_T f^2, 1\rangle = \langle f^2, W_T^* 1\rangle = \langle f^2, 1\rangle = \|f\|^2. \end{aligned}$$ □

The set $\mathcal{P}$ of polymorphisms obviously admits the natural structure of a semigroup with involution $*$. The set $\mathcal{P}$ is convex: for any two polymorphisms $F = (\varphi;\ p;\ I)$ and $G = (\psi;\ q;\ K)$ and $\alpha \in [0, 1]$ there is a polymorphism

$$\alpha F + (1 - \alpha)G := (\varphi, \psi;\ \alpha p, (1 - \alpha)q;\ I, K).$$

The boundary points of this family are naturally identified with G and F.

3.6. *Mixing and ergodicity.* Let $\Lambda \subset L_2$ be the space of W_T-invariant functions. Then by the von Neumann ergodic theorem for contracting operators [**3**] for any $f \in L_2$ the sequence

$$\frac{1}{N} \sum_{n=0}^{N-1} W_T^n f \tag{3.4}$$

converges in L_2 to $\pi(f)$, where $\pi : L_2 \to \Lambda$ is the orthogonal projection.

Definition 3.2. The polymorphism $T = (\varphi;\ p;\ I)$ is said to be ergodic if any fixed point of W_T is a constant: $\Lambda = \mathrm{span}(1)$.

The polymorphism T is said to be mixing if for any $f \in L_2$ we have

$$W_T^n f \to \overline{f} = \langle 1,\ f\rangle \text{ in the weak } L_2\text{-topology as } n \to \infty.$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

If T is ergodic, sequence (3.4) converges to $\pi(f) = \overline{f}$.

If T is mixing, it is ergodic. The inverse is not true: take for example $J = 1$,

$$T = (\varphi;\, 1;\, [0,\, 1]), \qquad \varphi(x) = x + \omega \bmod 1, \quad \omega = \text{const} \in \mathbb{R}\backslash\mathbb{Q}.$$

If T is mixing, T^* is also mixing. Indeed, $\lim_{n\to\infty}\langle W_T^n f,\, g\rangle = \overline{f}\,\overline{g}$ for any $f,\, g \in L_2$ if and only if for any $f,\, g \in L_2$ $\lim_{n\to\infty}\langle W_T^{*n} f,\, g\rangle = \overline{f}\,\overline{g}$.

Example. Consider the polymorphism $T = (\varphi_1,\, \varphi_2;\, 1/2,\, 1/2;\, [0,\, 1],\, [0,\, 1])$,

$$\varphi_1 = x/2, \quad \varphi_2 = x/2 + 1/2.$$

It is easy to check that

$$T^* = (\varphi;\, 1;\, [0,\, 1]), \quad \varphi(x) = 2x \bmod 1.$$

Then for any $f \in L_2$

$$W_T f(x) = \begin{cases} f(2x) & \text{if } x \in (0,\, 1/2), \\ f(2x-1) & \text{if } x \in (1/2,\, 1), \end{cases}$$

$$W_{T^*} f(x) = \frac{1}{2}\left(f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right)\right).$$

This implies (for example with the help of the Fourier expansions) that both T and T^* are mixing. The second statement is a standard fact from ergodic theory.

4. *The case* $I_1 = \cdots = I_J = [0,\, 1]$

A priori it is not obvious that conditions **P**, **M**, and **L** can be easily satisfied simultaneously. In this section we show that these conditions are quite flexible and the set of polymorphisms is sufficiently rich.

4.1. *General remarks.* Let $T = (\varphi;\, p;\, I)$ be a polymorphism. If all functions φ_j are increasing, T is said to be increasing.

Consider the case

$$I_1 = \cdots = I_J = \varphi_1(I_1) = \cdots = \varphi_J(I_J) = [0,\, 1].$$

For such polymorphisms we use the shorter notation $T = (\varphi;\, p)$. In this case $V(x) = V^*(y) = \{1,\, \ldots,\, J\}$ for all $x,\, y \in [0,\, 1]$.

If $T = (\varphi;\, p)$ is an increasing polymorphism, for all j we have $\varphi_j' \geq 0$, $\varphi_j(0) = 0$, and $\varphi_j(1) = 1$. In this case, for any $x \in [0,\, 1]$, (3.3) takes the form

$$\sum_{j=1}^{J} p_j \circ \psi_j(y)\psi_j'(y) = 1. \tag{4.1}$$

Putting $P_j(x) = \int_0^x p_j(\xi)\, d\xi$, we obtain

$$\sum_{j=1}^{J} P_j \circ \psi_j(y) = y. \tag{4.2}$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

4.2. *Construction of φ_J.*

PROPOSITION 4.1. *Given C^1-smooth $\varphi_1, \dots, \varphi_{J-1} : [0, 1] \to [0, 1]$, $\varphi_j' > 0$, $\varphi_j(0) = 0$, $\varphi_j(1) = 1$ and continuous $p_1, \dots, p_J : [0, 1] \to [0, 1]$, such that for some $c_1, c_2, c_3 \in [0, 1]$*

$$\sum_{j=1}^{J} p_j = 1, \quad \sum_{j=1}^{J-1} p_j \circ \psi_j \cdot \psi_j' < 1 - c_1,\ c_2 < p_J < c_3, \tag{4.3}$$

then there exists a unique $\varphi_J : [0, 1] \to [0, 1]$ such that $(\varphi;\ p)$ is a polymorphism, φ_J is C^1-smooth, $c_2 < \varphi_J' < c_3/c_2$, $\varphi_J(0) = 0$, and $\varphi_J(1) = 1$.

Proof. By (4.3) the function P_J^{-1} is C^1-smooth and $1/c_3 < (P_J^{-1})' < 1/c_2$. By using (4.2) we obtain

$$\psi_J(y) = P_J^{-1}\left(y - \sum_{j=1}^{J-1} P_j \circ \psi_j(y)\right). \tag{4.4}$$

According to (4.3),

$$c_1/c_3 < \psi_J' < 1/c_2.$$

We also have

$$\psi_J(0) = P_J^{-1}(0) = 0, \quad \psi_J(1) = P_J^{-1}\left(1 - \sum_{j=1}^{J-1} P_j(1)\right) = P_J^{-1}(P_J(1)) = 1. \qquad \square$$

PROPOSITION 4.2. *Let $(\varphi;\ p)$ be an increasing polymorphism. Suppose that $x \in [0, 1]$ is a point such that $\varphi_1(x) = \dots = \varphi_{J-1}(x) = x$. Then $\varphi_J(x) = x$.*

Indeed, by using (4.4) and (4.2), we obtain that $\psi_1(x) = \dots = \psi_{J-1}(x) = x$ implies that $\varphi_J(x) = x$. $\square$

4.3. *Construction of p.*

PROPOSITION 4.3. *Given smooth increasing functions $f, g : [0, 1] \to [0, 1]$ such that $f(0) = g(0) = 0$, $f(1) = g(1) = 1$, $f(x) < x < g(x)$ for $x \in (0, 1)$. Consider the function†*

$$\begin{aligned}\vartheta(x) &= f(x) - g^{-1} \circ f(x) + f \circ g^{-1} \circ f(x) - g^{-1} \circ f \circ g^{-1} \circ f(x) + \cdots \\ &= \sum_{n=1}^{+\infty} (g \circ (g^{-1} \circ f)^n(x) - (g^{-1} \circ f)^n(x)).\end{aligned} \tag{4.5}$$

Suppose that $\vartheta' \in (0, 1)$.

Then $(f, g;\ p, 1 - p)$ is an increasing polymorphism, where $p = \vartheta'$.

Proof. We put in (4.2) $J = 2$, $P_1(x) = P(x)$, $P_2(x) = x - P(x)$, $\psi_1 = f^{-1}$, $\psi_2 = g^{-1}$, and $y = f(x)$. Then we have

$$P(x) - P \circ g^{-1} \circ f(x) = f(x) - g^{-1} \circ f(x).$$

† Series (4.5) is alternating and absolute values of its terms monotonically tend to zero. Therefore it converges.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

Plugging in this equation $g^{-1}\circ f(x)$, $g^{-1}\circ f\circ g^{-1}\circ f(x), \dots$ instead of x and performing summation, we get

$$P(x) - P(0) = \vartheta(x).$$

Here we have used the fact that the sequence

$$f(x),\ g^{-1}\circ f(x),\ f\circ g^{-1}\circ f(x),\ g^{-1}\circ f\circ g^{-1}\circ f(x), \dots$$

tends to zero for any $x \in [0, 1]$. Hence $p(x) = P'(x) = \vartheta'(x)$. □

Now we present a condition necessary for the inclusion $\vartheta' \in [0, 1]$.

PROPOSITION 4.4. *Suppose that $\vartheta'(x)$ is bounded and*

$$\lim_{n\to+\infty} \frac{d}{dx}(g^{-1}\circ f)^{-n}(x) = 0 \quad \textit{for any } x \in (0, 1). \tag{4.6}$$

Then the function

$$\Theta(x) = \sum_{n=-\infty}^{+\infty} (g\circ(g^{-1}\circ f)^n(x) - (g^{-1}\circ f)^n(x))$$

does not depend on x.

Proof. For any $x \in (0, 1)$ we have

$$\lim_{n\to+\infty} \vartheta\circ(g^{-1}\circ f)^{-n}(x) = \Theta(x).$$

Moreover, by (4.6), we have the same equation for the derivatives:

$$\lim_{n\to+\infty} \frac{d}{dx}\,\vartheta\circ(g^{-1}\circ f)^{-n}(x) = \frac{d}{dx}\Theta(x).$$

Given any point $x_0 \in (0, 1)$ we put $x_n = (g^{-1}\circ f)^{-n}(x_0)$. Then

$$\frac{d}{dx}\bigg|_{x=x_0} \vartheta\circ(g^{-1}\circ f)^{-n}(x) = \vartheta'(x_k)\cdot\frac{d}{dx}(g^{-1}\circ f)^{-n}(x_0).$$

In the limit $n \to +\infty$ this quantity equals $\Theta'(x_0)$. By using (4.6) and boundedness of ϑ' we obtain $\Theta' = 0$. □

5. *Some non-ergodic polymorphisms*

A generic polymorphism should be expected to be ergodic and even mixing. However there are some exceptions. In this section we present two mechanisms which prevent ergodicity.

Let $T = (\varphi;\ p)$ be an increasing polymorphism such that for some point $x_0 \in (0, 1)$ $\varphi_1(x_0) = \dots = \varphi_{J-1}(x_0) = x_0$. Then, by Proposition 4.2, $\varphi_J(x_0) = x_0$ and the intervals $[0, x_0]$ and $[x_0, 1]$ are invariant. Such a polymorphism is not ergodic because $W_T f = f$ for any f equal to constants (may be distinct) on $(0, x_0)$ and $(x_0, 1)$.

This example is in a sense trivial because the dynamics, restricted to any of the intervals $(0, x_0)$ and $(x_0, 1)$ can be ergodic. Now we present a less obvious example of a 'non-ergodic', in fact, 'non-chaotic' polymorphism.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

PROPOSITION 5.1. *There exist* $f,\ p : [0,1] \to [0,1]$ *such that the following hold.*

(a) $f \in C^\infty([0,1])$.

(b) $f(x) < x$ *for any* $x \in (0,1)$.

(c) $f(0) = 0,\ f(1) = 1$.

(d) $f'|_{[0,1]} > 0$.

(e) $0 < p(x) < 1$ *for any* $x \in (0,1)$.

(f) $(f,\ f^{-1};\ p,\ 1-p)$ *is a polymorphism.*

Any increasing polymorphism $(f,\ f^{-1};\ p,\ 1-p)$ from Proposition 5.1 presents another mechanism of the absence of ergodicity. Indeed, let $[\alpha,\ \beta)$ be a 'fundamental' semi-interval, i.e.,

$$f^n([\alpha,\beta)) \cap f^m([\alpha,\beta)) = \emptyset \quad \text{for all integer } m \neq n, \text{ and } \bigcup_{n\in\mathbb{Z}} f^n([\alpha,\beta)) = (0,1).$$

For any $\rho_0 \in L_2$, supp $\rho_0 \subset [\alpha,\ \beta]$, obviously

$$\rho := \sum_{n=-\infty}^{\infty} \rho_0 \circ f^n \in L_2 \quad \text{and} \quad W_T \rho = \rho.$$

Proof of Proposition 5.1.

LEMMA 5.1. *Given* $\lambda > 0$ *there exists a function* $h : \mathbb{R} \to (0,1)$ *such that the following hold.*

(1) $h \in C^\infty(\mathbb{R})$.

(2) $h' > 0,\ h(-\infty) = 0,\ h(+\infty) = 1$.

(3) *The function* $h - 1/2$ *is odd.*

(4) $h'(y) \geq h'(y-1)$ *for* $y \leq 0$.

(5) $\sum_{n=-\infty}^{+\infty}(h(2n+y) - h(2n-1+y))$ *does not depend on* y.

(6) $h'(y-1)/h'(y) = e^\lambda(1 + O(1/y))$ *for* $y \to +\infty$.

Proposition 5.1 follows from Lemma 5.1. Indeed, for any $x \in (0,1)$ we put $f(x) = h(-1 + h^{-1}(x))$. For any natural n we obtain

$$f^n(x) = h(-n + h^{-1}(x)).$$

Then assertions (a), (b), and (c) obviously hold. We have to check (d) only for $x = 0$ and $x = 1$. For x close to 1,

$$f'(x) = \frac{h'(-1+y)}{h'(y)} = e^\lambda(1 + O(1/y)), \qquad y = h^{-1}(x).$$

Hence $\lim_{x\to 1} f'(x) = e^\lambda$. We have $\lim_{x\to 0} f'(x) = e^{-\lambda}$ because h' is odd.

To check (e), we note that, by condition (5),

$$\sum_{n=-\infty}^{+\infty} (h'(2n+y) - h'(2n-1+y)) = 0. \tag{5.1}$$

(The series converges according to condition (6).) As in Proposition 4.3 we define $p = \vartheta'$ (see (4.5)). The function ϑ in the case $g = f^{-1}$ has the form

$$\vartheta(x) = \sum_{n=1}^{+\infty}(f^{2n-1}(x) - f^{2n}(x)) = \sum_{n=1}^{+\infty}(h(1-2n+h^{-1}(x)) - h(-2n+h^{-1}(x))).$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

Therefore

$$\vartheta'(x) = \sum_{n=1}^{+\infty} \frac{h'(1-2n+y) - h'(-2n+y)}{h'(y)}, \quad y = h^{-1}(x).$$

Now we check (e). By Proposition 4.3, $p = \vartheta'$. Hence we have to prove the inequality

$$0 < \vartheta'(x) < 1 \quad \text{for } x \in (0, 1). \tag{5.2}$$

We put $y = h^{-1}(x)$. Then

$$\vartheta'(x) = \frac{h'(-1+y) - h'(-2+y) + h'(-3+y) - h'(-4+y) + \cdots}{h'(y)}.$$

First, suppose that $y \le 0$. The series in the numerator is alternating and by condition (4) absolute values of its terms monotonically decrease. Hence its sum S satisfies $0 < S < h'(-1+y)$. Since $h'(y) > h'(-1+y)$, we obtain (5.2).

Now suppose that $y = h^{-1}(x) > 0$. Then by using the fact that h' is odd, we have

$$\vartheta'(x) = \frac{h'(1-y) - h'(2-y) + h'(3-y) - h'(4-y) + \cdots}{h'(-y)}.$$

By (5.1), we get

$$\vartheta'(x) = \frac{h'(-y) - h'(-1-y) + h'(-2-y) - h'(-3-y) + \cdots}{h'(-y)}.$$

The series in the numerator is alternating and absolute values of its terms monotonically decrease. Hence its sum S satisfies $0 < S < h'(-y)$. This implies (5.2). □

We prove Lemma 5.1 in Appendix C.

6. *L_2-norm and W_T*

The main result of this section is Theorem 3 which shows that in some cases mixing property follows from ergodicity.

6.1. *On the difference $\|f\|^2 - \|W_T f\|^2$.* Let $T = (\varphi;\, p)$ be a polymorphism. Then by (3.1)

$$W_T f(x) = \sum_{j=1}^{J} q_j(x) f \circ \psi_j(x), \quad q_j(x) = p_j \circ \psi_j(x) |\psi_j'(x)|.$$

PROPOSITION 6.1. *For any $f \in L_2$,*

$$\|f\|^2 - \|W_T f\|^2 = \frac{1}{2} \sum_{j,k} B_{jk}(f), \tag{6.1}$$

$$B_{jk}(f) = \int_0^1 q_j(y) q_k(y) (f \circ \psi_j(y) - f \circ \psi_k(y))^2 \, dy. \tag{6.2}$$

COROLLARY 6.1. *Suppose that q_j, $j = 1, \ldots, J$, do not vanish almost everywhere. Then we have the following.*

(a) *The equation $\|W_T f\| = \|f\|$ implies that $f \circ \psi_j$ does not depend on j.*

(b) *The equation $W_T f = f$ implies that $f \circ \psi_j = f$ for all j.*

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

COROLLARY 6.2. *Suppose that q_j, $j = 1, \dots, J$, do not vanish almost everywhere. Then $T = (\varphi;\ p)$ is ergodic if and only if T^* is ergodic.*

Indeed, the equations $f \circ \psi_j = f$ and $f \circ \varphi_j = f$ are equivalent.

Let H_j, H_{jk} $(j \neq k)$ be the spaces defined as follows:

$$H_j = \{f \in L_2 \mid f = f \circ \psi_j\},$$
$$H_{jk} = \{f \in L_2 \mid f \circ \psi_j = f \circ \psi_k\}. \tag{6.3}$$

COROLLARY 6.3. *Suppose that q_j, $j = 1, \dots, J$ do not vanish almost everywhere and $\bigcap_{j=1}^{J} H_j = \mathrm{span}(1)$. Then T is ergodic.*

Proof of Proposition 6.1 is based on a direct calculation. By using the equations $\sum_j p_j = \sum_j q_j = 1$, we get

$$\|f\|^2 = \int \sum_j p_j(x) f^2(x)\, dx = \int \sum_j q_j(y) f^2 \circ \psi_j(y)\, dy$$
$$= \int \sum_{j,k} q_k(y) q_j(y) f^2 \circ \psi_j(y)\, dy,$$

$$\|W_T f\|^2 = \int \sum_{j,k} q_j(y) q_k(y) f \circ \psi_j(y)\ f \circ \psi_k(y)\, dy.$$

Equation (6.1) now follows. □

6.2. *A lower estimate for integral (6.2).* In this section we estimate integrals (6.2) in terms of $\|f\|$ provided f lies in $H_{jk}^{\perp}$. The spaces H_{jk}, defined in (6.3), can be presented in the form

$$H = \{f \in L_2 \mid f = f \circ h_{jk}\}, \quad h_{jk} = \psi_j \circ \varphi_k.$$

Below we fix a pair of indices $j \neq k$ and skip them for brevity: $H = H_{jk}$ and $h = h_{jk}$.

Consider a smooth or piecewise-smooth increasing function $h : [0, 1] \to [0, 1]$ such that the following hold.

(i) $h(0) = 0$, $h(1) = 1$, and $h'(x) > 0$ for any $x \in (0, 1)$ (if h is not differentiable at x, $h'(x)$ denotes right and left derivatives).

(ii) The function $h(x) - x$ does not have zeros on $(0, 1)$.

(iii) $\lim_{x \to 0} h'(x) = \lambda_0$, $\lim_{x \to 1} h'(x) = \lambda_1$, where $0 < \lambda_0 < 1 < \lambda_1$ or $0 < \lambda_1 < 1 < \lambda_0$.

We define the space

$$H = \{f \in L_2 \mid f = f \circ h\} \tag{6.4}$$

and its orthogonal complement $H^{\perp}$. For a positive function $Q : [0, 1] \to \mathbb{R}$, $Q > \hat{c} > 0$, consider the quadratic functional

$$B(f) = \int_0^1 Q(x)(f \circ h(x) - f(x))^2\, dx, \quad f \in L_2.$$

We put

$$c = \inf_{0 \neq f \in H^{\perp}} B(f)\|f\|^{-2} = \inf_{f \in H^{\perp}, \|f\| = 1} B(f). \tag{6.5}$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

LEMMA 6.1. *The constant c is positive.*

Proof of Lemma 6.1 is contained in Appendix B.

COROLLARY 6.4. *Suppose that $q_j, q_k > c_0 > 0$ and $\psi_j \circ \varphi_k$ satisfies conditions (i)–(iii). Then there exists a constant $c > 0$ such that for any $f \in L_2$*

$$\|f\| - \|W_T f\| \geq c\|\mathrm{pr}_{H_{jk}^{\perp}}(f)\|,$$

where $\mathrm{pr}_{H_{jk}^{\perp}} : L_2 \to H_{jk}^{\perp}$ is the orthogonal projection to $H_{jk}^{\perp}$.

6.3. *Mixing condition for $(\varphi;\ p)$.*

THEOREM 2. *Let $T = (\varphi_1, \ldots, \varphi_J;\ p_1, \ldots, p_J)$, $J \geq 3$ be a polymorphism such that $q_1, q_2, q_3 > c_0 > 0$ and the functions $\psi_1 \circ \varphi_2$, $\psi_1 \circ \varphi_3$ satisfy conditions (i)–(iii). Suppose that $H_{12} \cap H_{13} = \mathrm{span}(1)$. Then T is mixing.*

Proof. For any $f \in L_2$ we will prove that $W_T^n f$ weakly tends to $\langle 1,\ f\rangle$ as $n \to \infty$. It is sufficient to consider the case $\langle 1,\ f\rangle = 0$ because

$$W_T(1) = 1 \quad \text{and} \quad W_T(\mathrm{span}(1)^{\perp}) \subset \mathrm{span}(1)^{\perp}.$$

We write $\mathcal{H}_{12} = H_{12} \cap \mathrm{span}(1)^{\perp}$, $\mathcal{H}_{13} = H_{13} \cap \mathrm{span}(1)^{\perp}$. Then $\mathcal{H}_{12} \cap \mathcal{H}_{13} = 0$.

Let H be a Hilbert space and Π its closed subspace. For any non-zero $v \in H$ we define the angle between v and Π:

$$\widehat{(v,\ \Pi)} = \arccos\left(\sup_{0 \neq w \in \Pi} \frac{\langle v,\ w\rangle}{\|v\|\ \|w\|}\right) = \arcsin\left(\inf_{0 \neq w \in \Pi^{\perp}} \frac{\langle v,\ w\rangle}{\|v\|\ \|w\|}\right).$$

Consider the sequence

$$f_0,\ f_1, \ldots \in \mathrm{span}(1)^{\perp}, \qquad f_k = W_T^k f.$$

If at least one of the sequences $(\widehat{f_k,\ \mathcal{H}_{12}})$ or $(\widehat{f_k,\ \mathcal{H}_{13}})$ does not tend to zero, by Corollary 6.4 we have $\|f_k\| \to 0$. Hence it remains to consider the case when $\lim_{k\to\infty} (\widehat{f_k,\ \mathcal{H}_{12}}) = \lim_{k\to\infty} (\widehat{f_k,\ \mathcal{H}_{12}}) = 0$. In this case the proof directly follows from the lemma.

LEMMA 6.2. *Let $\Pi_1, \Pi_2 \subset H$, $\Pi_1 \cap \Pi_2 = 0$ be two closed subspaces in the Hilbert space H. Suppose that the sequence $v_1, v_2, \ldots \in B_c$ is such that $\|v_k\| \leq c$, $c > 0$, and*

$$\lim_{k\to\infty} (\widehat{v_k,\ \Pi_1}) = \lim_{k\to\infty} (\widehat{v_k,\ \Pi_2}) = 0.$$

Then $v_k \to 0$ weakly as $k \to \infty$.

Proof of Lemma 6.2. For any $w \in \Pi_1^{\perp}$ we have

$$\lim_{j\to\infty} \frac{\langle v_j,\ w\rangle}{\|v_j\|} = 0.$$

Since $\|v_j\| \leq c$, this implies that

$$\lim_{j\to\infty} \langle v_j,\ w\rangle = 0. \tag{6.6}$$

Analogously, equation (6.6) holds for any $w \in \Pi_2^{\perp}$.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

Since $\Pi_1 \cap \Pi_2 = 0$ and Π_1, Π_2 are closed, the space $\Pi_1^\perp \oplus \Pi_2^\perp$ is dense in H. (Indeed, if $\overline{\Pi_1^\perp \oplus \Pi_2^\perp} \neq H$, we have $\emptyset \neq \overline{\Pi_1^\perp \oplus \Pi_2^\perp}^\perp \subset \Pi_1 \cap \Pi_2$.) Hence (6.6) holds for any $w \in H$. □

THEOREM 3. *Let $T = (\varphi;\ p)$ be an ergodic polymorphism, $q_1, \dots, q_J > c_0 > 0$ and the functions $\psi_j \circ \varphi_k$ satisfy conditions (i)–(iii). Then T is mixing.*

Proof. For any $f \in L_2$ we will show that the sequence $f_k = W_T^k f$ weakly tends to $\langle 1,\ f\rangle$. It is sufficient to consider the case $\langle 1,\ f\rangle = 0$.

Consider an arbitrary weakly converging subsequence $f_{k_l} \to \hat{f}$. Then for any $j, s \in \{1, \dots, J\}$ and $w \in H_{js}^\perp$. we have $\langle f_{k_l}, w\rangle \to \langle \hat{f}, w\rangle$. The last quantity is always zero (otherwise by Corollary 6.4 we have $\|f_{k_l}\| \to 0$). Hence, $\hat{f} \in H_{js}$, i.e., $\hat{f} \circ \psi_1 = \cdots = \hat{f} \circ \psi_J$.

Since $\hat{f} = W_T \hat{f}$ we obtain

$$\hat{f} = \sum q_j \hat{f} \circ \psi_j = \hat{f} \circ \psi_1 \sum q_j = \hat{f} \circ \psi_1.$$

Analogously $\hat{f} = \hat{f} \circ \psi_j$, $j = 2, \dots, J$. By using ergodicity of T and the equation $\langle 1,\ \hat{f}\rangle = 0$, we get: $\hat{f} = 0$.

Thus the limit of any weakly converging subsequence of f_k is zero. This implies that f_k weakly converges to zero. Indeed, suppose that there exists $w \in L_2$ and a subsequence f_{k_s} such that

$$|\langle f_{k_s}, w\rangle| > \hat{c} > 0. \tag{6.7}$$

Since the norms $\|f_k\|$ are bounded, we can choose in f_{k_s} a weakly converging subsequence. We obtain a contradiction with (6.7) because its weak limit should be zero. □

7. *Polymorphisms $T_{\beta,s}$*

In this section we study one two-parametric family of polymorphisms. The main result is that for an open set in the parameter plane these polymorphisms are mixing.

Consider the polymorphism $T_{\beta,s} = (\varphi_1, \varphi_2;\ 1/2,\ 1/2)$, where the functions $\varphi_{1,2}$: $[0,\ 1] \to [0,\ 1]$ are defined in (2.5)–(2.6),

$$s, \beta \in (0,\ 1), \quad \alpha + \beta = 2, \alpha s < 1. \tag{7.1}$$

Graphs of the functions φ_1 and φ_2 are presented in Figure 2. Note that the case $\beta = 0$, $s = 1/2$ was considered in §3.6.

Below we assume that α, β, and s satisfy the following condition.

C For some $l, m \in \mathbb{N}$,

$$\frac{\alpha - s}{1 - \beta s} < \beta^l \alpha^m < \frac{1}{\beta}, \tag{7.2}$$

$$\frac{\alpha}{1 + s(1-\beta)} < \beta^l \alpha^m < \frac{1}{\beta} \frac{\alpha}{1 + s(1-\beta)}, \tag{7.3}$$

$$\beta^l \alpha^m < \frac{1 - s}{1 - \beta} \frac{\alpha}{1 + s(1-\beta)}. \tag{7.4}$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

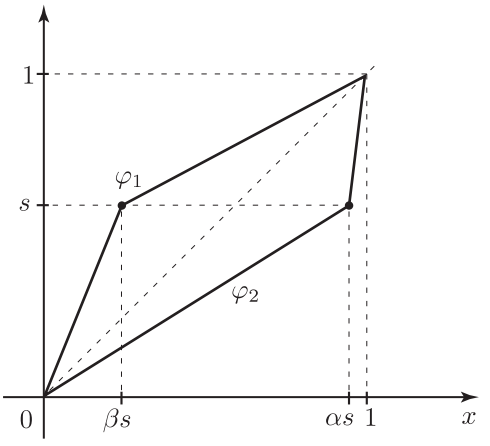


FIGURE 2. Polymorphism $T_{\beta,s}$.

If $\beta \le s$, the inequalities are incompatible because the right-hand side of (7.4) is not greater than the left-hand side of (7.3). If $\beta > s$, inequalities (7.2)–(7.4) hold for $l = 0$, $m = 1$; however, **C** requires $l, m > 0$.

THEOREM 4. *Suppose that condition* **C** *holds. Then the polymorphism $T_{\beta,s}$ is mixing.*

Remarks. (1) The change of variable $x \mapsto 1 - x$ transforms $T_{\beta,s}$ to $T_{\beta',s'}$, where

$$\beta' = \frac{1 - \alpha s}{1 - s}, \quad s' = 1 - s.$$

Hence Theorem 4 remains true if in condition **C** we replace α, β, s by α', β', s'. Note that the inequalities $\beta > s$ and $\beta' > s'$ are equivalent.

(2) Condition **C** obviously holds if $\beta > s$ and $\log \alpha / \log \beta$ is irrational, because in this case the numbers $\beta^l \alpha^m$, $l, m \in \mathbb{N}$, form a dense set on $[0, \infty)$. Hence the set of β, s, for which condition **C** holds, is open and dense in the two-dimensional domain

$$\{(\beta, s) : \beta > s \text{ and conditions (7.1) hold}\}.$$

(3) We expect that condition **C** is technical and Theorem 4 holds for all β, s, satisfying (7.1).

Proof of Theorem 4. We will prove that $T_{\beta,s}$ is ergodic. Then Theorem 3 implies that it is mixing.

Let $f \in L_2$ satisfy the equation $W_{T_{\beta,s}} f = f$. Then by Corollary 6.1(b) we have

$$f = f \circ \varphi_1 = f \circ \varphi_2. \tag{7.5}$$

We will prove that (7.5) implies $f =$ constant.

The further argument is based on the following lemma. Consider the interval $(a, b) \in \mathbb{R}$ and two linear contractions $h_a, h_b : (a, b) \to (a, b)$,

$$h_a(x) = \lambda_a(x - a) + a, \quad h_b(x) = \lambda_b(x - b) + b, \; 0 < \lambda_a, \lambda_b < 1.$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

LEMMA 7.1. *Suppose that* $\lambda_a + \lambda_b > 1$. *Then any measurable function* $g : (a, b) \to \mathbb{R}$, *satisfying the equations* $g = g \circ h_a = g \circ h_b$ *equals a constant almost everywhere.*

We prove the lemma in Appendix D and now return to the proof of Theorem 4. To apply Lemma 7.1 we construct two maps Φ_a, Φ_b such that the following hold.

(a) Φ_a and Φ_b are finite compositions of φ_1 and φ_2.

(b) For some intervals I_a, $I_b \subset (0, 1)$ the restrictions $\hat{\Phi}_a = \Phi_a|_{I_a}$ and $\hat{\Phi}_b = \Phi_b|_{I_b}$ are linear contractions.

(c) $\Phi_a|_{(a,b)} = h_a$ and $\Phi_b|_{(a,b)} = h_b$ for some points $a, b \in I_a \cap I_b$, where $0 < \lambda_a$, $\lambda_b < 1$ and $\lambda_a + \lambda_b > 1$.

(d) $\bigcup_{k=-\infty}^{\infty} \varphi_1^k(I_b) = (0, 1)$.

Conditions (a)–(c) imply that the maps $h_a = \Phi_a|_{(a,b)}$ and $h_b = \Phi_b|_{(a,b)}$ satisfy Lemma 7.1. Therefore any f, satisfying (7.5), is a constant almost everywhere on (a, b) and, moreover, on $I_a \cup I_b$. By (d), the condition $f|_{I_b} = \text{constant}$ implies that $f = \text{constant}$ almost everywhere on $[0, 1]$.

Below it can happen that $b < a$. However we still denote the corresponding interval (a, b). We put

$$\Phi_b = \varphi_2 \circ \varphi_1 = \frac{1}{\alpha}\left(\frac{1-s}{1-\beta s}x + \frac{s-\beta s}{1-\beta s}\right), \quad I_b = \left(\beta s, \frac{1-\alpha\beta s}{1-s}s\right).$$

Since $s = \varphi_1(\beta s)$ and $\alpha\beta < 1$, we have $(\beta s, s) \subset I_b$. Therefore condition (d) is satisfied.

The map $\hat{\Phi}_b$ is a linear contraction with

$$\lambda_b = \frac{1-s}{\alpha(1-\beta s)}, \quad b = \frac{s}{1+s(1-\beta)} \in I_b.$$

Proof of the inequality $\beta s < b < (1 - \alpha\beta s)/(1 - s)s$ is straightforward.

For $l, m \in \mathbb{N}$ satisfying (7.2)–(7.4) we put

$$\Phi_a = \varphi_1^l \circ \varphi_2^m \circ \varphi_1.$$

Below we show that

$$I_a \subset I_b \quad \text{and} \quad \Phi_a(I_a) \subset (0, s). \tag{7.6}$$

Inclusions (7.6) imply

$$\Phi_a = \frac{\alpha}{\omega}\Phi_b = \frac{1}{\omega}\left(\frac{1-s}{1-\beta s}x + \frac{s-\beta s}{1-\beta s}\right), \quad \omega = \beta^l\alpha^m. \tag{7.7}$$

We put

$$a = \frac{(1-\beta)s}{\omega(1-\beta s) - 1 + s}.$$

Then $a \in (\beta s, s)$ because this inclusion is equivalent to (7.2). By (7.7) we obtain $\Phi_a(a) = a$.

The inclusion $\Phi_a(b) \in (\beta s, s)$ is equivalent to (7.3). We take I_a equal to a small neighborhood of (a, b). Then $I_a \subset I_b$ and $\Phi_a(I_a) \subset (\beta s, s)$. Therefore conditions (7.6) are satisfied.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

Now conditions (a) and (b) also hold. The map $\hat{\Phi}_b$ is a linear contraction with

$$\lambda_a = \frac{1-s}{\omega(1-\beta s)}.$$

Finally let us turn to condition (c). Proof of the inequality $0 < \lambda_b < 1$ is straightforward. The condition $0 < \lambda_a < 1$ is equivalent to the inequality $\omega > (1-s)/(1-\beta s)$ which follows from (7.2). The condition $\lambda_a + \lambda_b > 1$ is equivalent to (7.4).

Acknowledgements. The authors thank G. Piftankin and O. Zubelevich for useful discussions. The work is partially supported by RFBR grants (projects 08-01-00681-a, 09-01-00333-a) and the Program for Supporting Leading Scientific Schools (Grant NSh-1312.2006.1).

A. *Appendix. Preservation of the measure*

In this section we prove Theorem 1. It is sufficient to prove that any map $T_{\tau_s}^{\tau_{s+1}}$ pulls back the measure du on $U(\tau_{s+1})$ to the measure du on $U(\tau_s)$. We can assume that (τ_s, τ_{s+1}) does not contain critical points of A_0, A_+, and A_-.

Now we will have to consider several cases, depending on signs of A_0', A_+', and A_-' on (τ_s, τ_{s+1}):

$$\begin{aligned}
(+++) &\quad A_0' > 0 \quad \text{and} \quad A_\pm' > 0,\\
(+\pm) &\quad A_0' > 0 \quad \text{and} \quad A_+'A_-' < 0,\\
(---) &\quad A_0' < 0 \quad \text{and} \quad A_\pm' < 0,\\
(-\pm) &\quad A_0' < 0 \quad \text{and} \quad A_+'A_-' < 0.
\end{aligned}$$

We will discuss in details the cases $(+++)$ and $(+\pm)$, $A_+' < 0$. The cases $(---)$ and $(+\pm)$, $A_+' > 0$ are completely similar to the case $(+\pm)$, $A_+' < 0$. The case $(-\pm)$ is completely similar to the case $(+++)$.

Case $(+++)$. In this case $T_{\tau_s}^{\tau_{s+1}}$ equals the identity on

$$(0, A_+(\tau_s)) \sqcup (0, A_-(\tau_s)) \sqcup (A_0(\tau_{s+1}), C).$$

Hence we just need to consider the restriction $\mathcal{T}_s = T_{\tau_s}^{\tau_{s+1}}|_{(A_0(\tau_s), A_0(\tau_{s+1}))}$

$$\mathcal{T}_s : (A_0(\tau_s), A_0(\tau_{s+1})) \to (A_+(\tau_s), A_+(\tau_{s+1})) \sqcup (A_-(\tau_s), A_-(\tau_{s+1})).$$

Let u_0 be an arbitrary point in $(A_0(\tau_s), A_0(\tau_{s+1}))$. It has then two images: $u_\pm \in (A_\pm(\tau_s), A_\pm(\tau_{s+1}))$. By construction of $T_{\tau_s}^{\tau_{s+1}}$ (see §2, (b_0)), for some $\tilde{\tau} \in (\tau_s, \tau_{s+1})$, $u_0 = A_0(\tilde{\tau})$,

$$u_\pm = A_\pm(\tilde{\tau})$$

and u_0 is mapped to $u_\pm$ with probability $p_\pm$,

$$p_\pm = p_\pm(\tilde{\tau}) = \frac{A_\pm'(\tilde{\tau})}{A_0'(\tilde{\tau})}.$$

Hence $p_\pm/(du_\pm/du_0) = 1$. This equation coincides with (2.4) in case $(+++)$.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

Case $(+\pm)$, $A'_+ < 0$. In this case $T_{\tau_s}^{\tau_{s+1}}$ equals the identity on

$$(A_0(\tau_{s+1}),\, C) \sqcup (0,\, A_+(\tau_{s+1})) \sqcup (0,\, A_-(\tau_s)).$$

We only need to consider $\hat{T}_s = T_{\tau_s}^{\tau_{s+1}}|_{(A_0(\tau_s),A_0(\tau_{s+1}))\sqcup(A_+(\tau_{s+1}),A_+(\tau_s))}$

$$\hat{T}_s \,:\, (A_0(\tau_s),\, A_0(\tau_{s+1})) \sqcup (A_+(\tau_{s+1}),\, A_+(\tau_s)) \to (A_-(\tau_s),\, A_-(\tau_{s+1})).$$

Let u_- be an arbitrary point in $(A_-(\tau_s),\, A_-(\tau_{s+1}))$. It has two preimages: $u_0 \in (A_0(\tau_s),\, A_0(\tau_{s+1}))$ and $u_+ \in (A_+(\tau_{s+1}),\, A_+(\tau_s))$. By construction of $T_{\tau_s}^{\tau_{s+1}}$, for some $\tilde{\tau} \in (\tau_s,\, \tau_{s+1})$,

$$u_- = A_-(\tilde{\tau}), \quad u_0 = A_0(\tilde{\tau}), \quad u_+ = A_+(\tilde{\tau}),$$

where u_0 and u_+ are mapped to u_- with probability 1. Hence

$$\frac{1}{|du_-/du_0|} + \frac{1}{|du_-/du_+|} = \frac{A'_0(\tilde{\tau})}{A'_-(\tilde{\tau})} - \frac{A'_+(\tilde{\tau})}{A'_-(\tilde{\tau})} = 1.$$

This equation coincides with (2.4) in the case $(+\pm)$, $A'_+ < 0$. □

B. *Appendix. Lower estimate for the functional B*

B.1. *Proof of Lemma 6.1.* Since Q is positive, it is possible to assume that $Q \equiv 1$. Below we fix a point $x_0 \in (0,\, 1)$ and put $x_1 = h(x_0)$. We have the equations

$$\|f\|^2 = \int_{x_0}^{x_1} \sum_{-\infty}^{\infty} f^2 \circ h^k(x)(h^k)'(x)\, dx,$$

$$B(f) = \int_{x_0}^{x_1} \sum_{-\infty}^{\infty} (f \circ h^{k+1}(x) - f \circ h^k(x))^2 (h^k)'(x)\, dx. \tag{B.1}$$

Let $\mathrm{pr}_H : L_2 \to H$ be the orthogonal projection. Then, for any $f \in H^\perp$,

$$\mathrm{pr}_H f = \frac{\sum_{k\in\mathbb{Z}} (h^k)' f \circ h^k}{\sum_{k\in\mathbb{Z}} (h^k)'} = 0. \tag{B.2}$$

For any $x \in [x_0,\, x_1]$ we put $\chi_k = \chi_k(x) = (h^k)'(x)$. For definiteness we assume that in (iii) (§6.2) $0 < \lambda_1 < 1 < \lambda_0$. Then, by (i)–(iii), we have the following.

(1) $\chi_k > 0$.

(2) $\lim_{k\to-\infty} \chi_k/\lambda_0^k$ and $\lim_{k\to+\infty} \chi_k/\lambda_1^k$ exist and are positive (i.e., $\chi_k(x)$, $x \in [x_0,\, x_1]$ exponentially tends to zero as $k \to \pm\infty$).

(3) There exist $c_1,\, c_2 > 0$ such that

$$c_1 \chi_k(\xi) \le \chi_k(x) \le c_2 \chi_k(\xi) \quad \text{for any } k \in \mathbb{Z} \text{ and any } x,\, \xi \in [x_0,\, x_1].$$

Conditions (1)–(3) and equation (B.2) give us the possibility to apply Lemmas B.1–B.2 (see Appendix B.2). For any $x \in [x_0,\, x_1]$ these lemmas estimate the integrand in (B.1) provided $f \in H^\perp$. Let $\gamma = \gamma(\chi(x))$ be defined by (B.3). By Lemmas B.1 and B.2, $\gamma(\chi(x)) > \hat{\gamma} > 0$. We have

$$B(f) \ge \int_{x_0}^{x_1} \sum_{k=-\infty}^{+\infty} \gamma(\chi(x)) f^2 \circ h^k(x)(h^k)'(x)\, dx$$

$$\ge \hat{\gamma} \int_{x_0}^{x_1} \sum_{k=-\infty}^{+\infty} f^2 \circ h^k(x)(h^k)'(x)\, dx = \hat{\gamma}\|f\|^2.$$

Hence in (6.5) we can take $c = \hat{\gamma}$. □

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

B.2. *Discrete version of Lemma 6.1.* Consider the sequence of positive numbers $\dots, \chi_{-1}, \chi_0, \chi_1, \dots$ such that for some constants $0 < \lambda_1 < 1 < \lambda_0$ and $\alpha_0, \alpha_1 > 0$

$$\lim_{n\to-\infty} \chi_n/\lambda_0^n = \alpha_0, \quad \lim_{n\to+\infty} \chi_n/\lambda_1^n = \alpha_1.$$

Consider the real Hilbert space $l_2(\chi)$, formed by sequences $a = (\dots, a_0, a_1, \dots)$. The scalar product and the norm are determined by

$$\langle a, b\rangle = \sum_{n\in\mathbb{Z}} \chi_n a_n b_n, \quad \|a\| = \sqrt{\langle a, a\rangle}.$$

We put $1 = (\dots, 1, 1, \dots) \in l_2(\chi)$,

$$\mathcal{B}(a) = \sum_{n\in\mathbb{Z}} \chi_n (a_{n+1} - a_n)^2,$$

$$\gamma = \inf_{a\neq 0, \langle a, 1\rangle = 0} \mathcal{B}(a)\|a\|^{-2} = \inf_{\|a\|=1, \langle a, 1\rangle = 0} \mathcal{B}(a). \tag{B.3}$$

LEMMA B.1. *The constant $\gamma = \gamma(\chi)$ is positive.*

LEMMA B.2. *Suppose that the sequences χ' and χ'' are such that for some positive constants $c_1 < c_2$*

$$c_1\chi_n'' \le \chi_n' \le c_2\chi_n'', \quad n \in \mathbb{Z}. \tag{B.4}$$

Then

$$\frac{c_1}{c_2}\gamma(\chi'') \le \gamma(\chi') \le \frac{c_2}{c_1}\gamma(\chi'').$$

Proof of Lemma B.1. For any $N \in \mathbb{N}$ we put

$$\langle a, b\rangle_N = \sum_{n=-N}^{N} \chi_n a_n b_n, \quad \|a\|_N = \sqrt{\langle a, a\rangle_N}, \quad \mathcal{B}_N(a) = \sum_{n=-N}^{N-1} \chi_n (a_{n+1} - a_n)^2,$$

$$\langle a, b\rangle_N^{\pm} = \sum_{\pm n > N} \chi_n a_n b_n, \quad \|a\|_N^{\pm} = \sqrt{\langle a, a\rangle_N^{\pm}},$$

$$\mathcal{B}_N^+(a) = \sum_{n\ge N} \chi_n (a_{n+1} - a_n)^2, \quad \mathcal{B}_N^-(a) = \sum_{n<-N} \chi_n (a_{n+1} - a_n)^2.$$

Then for all $a \in l_2(\chi)$ and $N \in \mathbb{N}$

$$\|a\|^2 = \|a\|_N^2 + (\|a\|_N^+)^2 + (\|a\|_N^-)^2, \quad \mathcal{B}(a) = B_N(a) + B_N^+(a) + B_N^-(a).$$

Let $a^{(1)}, a^{(2)}, \dots \in l_2(\chi)$ be a sequence such that

$$\langle a^{(j)}, 1\rangle = 0, \quad \|a^{(j)}\| = 1 \quad \text{and} \quad \lim_{j\to\infty} \mathcal{B}(a^{(j)}) = \gamma.$$

(1) Suppose that there exist $N, \beta > 0$ such that $\|a^{(j)}\|_N \ge \beta$ for any $j \in \mathbb{N}$. Then

$$|\langle a^{(j)}, 1\rangle_N| = \left|\langle a^{(j)}, 1\rangle_N^+ + \langle a^{(j)}, 1\rangle_N^-\right| \le \|a^{(j)}\|_N^+\|1\|_N^+ + \|a^{(j)}\|_N^-\|1\|_N^-$$
$$\le \sqrt{1-\beta^2}(\|1\|_N^+ + \|1\|_N^-).$$

Obviously $\|1\|_N^{\pm}$ tend to zero as $N \to \infty$. Therefore (since we can assume that N is sufficiently large)

$$|\langle a^{(j)}, 1\rangle_N| \le \tfrac{1}{2}\beta\|1\|_N.$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

Then for any j

$$\mathcal{B}(a^{(j)}) \geq \mathcal{B}_N(a^{(j)}) \geq \inf_{\|a\|_N=\beta, |\langle a, I\rangle_N| \leq (1/2)\beta\|I\|_N} \mathcal{B}_N(a)$$
$$= \beta^2 \inf_{\|a\|_N=1, |\langle a, I\rangle_N| \leq (1/2)\|I\|_N} \mathcal{B}_N(a).$$

For any $N > 0$ the last infimum is positive because it is, in fact, a minimum of a strictly positive continuous function of $2N + 1$ variables on a compact set.

(2) Suppose that there exists a subsequence j_s such that $\|a^{(j_s)}\|_N \to 0$ as $s \to \infty$ for any $N > 0$. We assume that N is so large that

$$|\chi_{-n}/\lambda_0^{-n} - \alpha_0| < \varepsilon_1\alpha_0, \quad |\chi_n/\lambda_1^n - \alpha_1| < \varepsilon_1\alpha_1 \text{ for any } n \geq N, \tag{B.5}$$

where $\varepsilon_1 > 0$ can be chosen arbitrarily small. Consider $a = a^{(j_s)}$ such that

$$\|a\|_{N+1} < \varepsilon_2, \tag{B.6}$$
$$\mathcal{B}(a) < \gamma + \varepsilon_3. \tag{B.7}$$

Then

$$(\|a\|_{N+1}^+)^2 + (\|a\|_{N+1}^-)^2 > 1 - \varepsilon_2^2.$$

One of the two terms in the left-hand side of this inequality is greater than $(1 - \varepsilon_2^2)/2$. For definiteness, we assume that

$$(\|a\|_{N+1}^+)^2 > (1 - \varepsilon_2^2)/2. \tag{B.8}$$

For any $a \in l_2(\chi)$ we define $\vec{a} \in l_2(\chi)$ such that $\vec{a}_n = a_{n+1}$ for all $n \in \mathbb{Z}$. Then

$$B_N^+(a) = (\|a\|_N^+)^2 + (\|\vec{a}\|_N^+)^2 - 2\langle a, \vec{a}\rangle_N^+ \geq (\|a\|_N^+ - \|\vec{a}\|_N^+)^2. \tag{B.9}$$

PROPOSITION B.1. *Inequalities (B.5)–(B.6) imply*

$$(\|a\|_N^+)^2 = \lambda_1(\|\vec{a}\|_N^+)^2 + \mathcal{R},$$

where

$$|\mathcal{R}| \leq \varepsilon_2^2 + \frac{2\varepsilon_1}{1 - \varepsilon_1}(\|a\|_{N+1}^+)^2. \tag{B.10}$$

Proof of Proposition B.1. By definition

$$(\|a\|_N^+)^2 = \sum_{n=N+1}^{\infty} \chi_n a_n^2 = \lambda_1(\|\vec{a}\|_N^+)^2 + \mathcal{R},$$
$$\mathcal{R} = \chi_{N+1}a_{N+1}^2 + \sum_{n=N+2}^{\infty} (\chi_n - \lambda_1\chi_{n-1})a_n^2.$$

Note that by (B.5)

$$|\chi_n - \lambda_1\chi_{n-1}| = \lambda_1^n\left|(\chi_n\lambda_1^{-n} - \alpha_1) - (\chi_{n-1}\lambda_1^{1-n} - \alpha_1)\right| \leq 2\varepsilon_1\lambda_1^n\alpha_1.$$

Therefore again by (B.5)

$$\sum_{n=N+2}^{\infty} (\chi_n - \lambda_1\chi_{n-1})a_n^2 \leq \sum_{n=N+2}^{\infty} 2\varepsilon_1\alpha_1\lambda_1^n a_n^2$$
$$\leq \sum_{n=N+2}^{\infty} \frac{2\varepsilon_1\chi_n a_n^2}{1 - \varepsilon_1} = \frac{2\varepsilon_1}{1 - \varepsilon_1}(\|a\|_{N+1}^+)^2. \tag{B.11}$$

Inequalities (B.6) and (B.11) imply estimate (B.10). □

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

Proposition B.1 and estimates (B.8) and (B.9) imply that

$$B_N^+(a) \geq (\|a\|_N^+ - \|\vec{a}\|_N^+)^2 = \frac{((\|a\|_N^+)^2 - (\|\vec{a}\|_N^+)^2)^2}{(\|a\|_N^+ + \|\vec{a}\|_N^+)^2}$$
$$\geq \frac{1}{4}\left(\left(\frac{1}{\lambda_1} - 1\right)(\|a\|_N^+)^2 - \frac{|R|}{\lambda_1}\right)^2$$
$$\geq \frac{1}{4}\left(\left(\frac{1}{\lambda_1} - 1 - \frac{2\varepsilon_1}{\lambda_1(1-\varepsilon_1)}\right)\frac{1-\varepsilon_2^2}{2} - \varepsilon_2^2\right)^2.$$

It follows from (B.7) that the last expression is smaller than $\gamma + \varepsilon_3$. If ε_1, ε_2, and ε_3 are sufficiently small, this means that $\gamma > (1-\lambda_1)^2/16\lambda_1^2 > 0$.

The case, when instead of (B.8) we have $(\|a\|_{N+1}^-)^2 > (1-\varepsilon_2^2)/2$, is analogous. □

Proof of Lemma B.2. The vector spaces $l_2(\chi')$ and $l_2(\chi'')$ coincide. Let $\|\cdot\|'$ and $\|\cdot\|''$ be the corresponding norms. Analogously we define $\mathcal{B}'$ and $\mathcal{B}''$. Now estimates (B.4) follow from the inequalities

$$\alpha_1\|a\|'' \leq \|a\|' \leq \alpha_2\|a\|'', \quad \alpha_1\mathcal{B}''(a) \leq \mathcal{B}'(a) \leq \alpha_2\mathcal{B}''(a).$$ □

C. *Appendix. Proof of Lemma 5.1*

Take any sequence $\{a_n\}$, $n \in \mathbb{Z}$ such that the following hold.

(i) $a_n + a_{-n} = 1$ for all $n \in \mathbb{Z}$.
(ii) $a_n < a_{n+1}$ for any $n \in \mathbb{Z}$.
(iii) The sequence $a_n - a_{n+1}$, $n = 0, 1, \ldots$ is decreasing.
(iv) $\lim_{n\to-\infty} a_n = 0$, $\lim_{n\to+\infty} a_n = 1$.
(v) $\sum_{n=-\infty}^{+\infty}(a_{2n} - a_{2n-1}) = 1/2$.
(vi) $(a_n - a_{n-1})/(a_{n-1} - a_{n-2}) = e^{\lambda}(1 + O(1/n))$ for $n \to -\infty$.

For example, we can take

$$a_n = \frac{1}{\sinh(\lambda/2)} \sum_{k=-\infty}^{n} e^{-\lambda|k-1/2|}.$$

Consider the function $h_0 : \mathbb{R} \to (0, 1)$ such that $h_0(n) = a_n$ for $n \in \mathbb{Z}$ and h_0 is linear on any interval $(n, n+1)$. Then h_0 satisfies conditions (2) and (3) of the lemma.

Instead of condition (4), h_0 satisfies the following condition:

$$h_0'(y) \geq h_0'(y+1) \quad \text{for any } y > -1,\ y \notin \mathbb{Z}. \tag{C.1}$$

The function $\chi(y) = \sum_{n=-\infty}^{+\infty}(h_0(2n+y) - h_0(2n-1+y))$ is 2-periodic in $y \in \mathbb{R}$, piecewise linear, and smooth outside $\mathbb{Z}$. By (v) $\chi(0) = 1/2$. Since $\chi(0) + \chi(1) = 1$, we have $\chi(1) = 1/2$. Therefore $\chi \equiv 1/2$ and condition (5) also holds.

Condition (vi) implies that

$$h_0'(y)/h_0'(y-1) = e^{\lambda}(1 + O(1/y)) \quad \text{for } y \to -\infty,\ y \notin \mathbb{Z}. \tag{C.2}$$

To satisfy conditions (1), (4), and (6), we perform a smoothing of h_0. To this end we consider a smooth even function $\omega : \mathbb{R} \to [0, +\infty)$ such that

$$\omega(\xi) = 0 \quad \text{for } |\xi| \geq 1 \text{ and } \int_{-\infty}^{+\infty} \omega(\xi)\, d\xi = 1.$$

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

We put

$$h(y) = \int_{-\infty}^{+\infty} h_0(\xi)\omega(y-\xi)\,d\xi \equiv \int_{-\infty}^{+\infty} h_0(y-\xi)\omega(\xi)\,d\xi.$$

The function h satisfies conditions (1)–(3). For any $y \geq 0$, (C.1) implies that

$$h'(y) - h'(y+1) = \int_{-1}^{1} (h_0'(y-\xi) - h_0'(y+1-\xi))\omega(\xi)\,d\xi \geq 0,$$

i.e., condition (4) holds. Condition (5) follows from the equation

$$\begin{aligned}
&\sum_{n=-\infty}^{+\infty} (h(2n+y) - h(2n-1+y)) \\
&\quad = \int_{-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} (h_0(2n+y-\xi) - h_0(2n-1+y-\xi))\omega(\xi)\,d\xi \\
&\quad = \int_{-\infty}^{+\infty} \frac{1}{2}\omega(\xi)\,d\xi = \frac{1}{2}.
\end{aligned}$$

By using (C.2) we obtain

$$\begin{aligned}
h'(y) &= \int_{-1}^{1} h_0'(y-\xi)\omega(\xi)\,d\xi = \int_{-1}^{1} e^{\lambda}(1+o(1/y))h_0'(-1+y-\xi)\omega(\xi)\,d\xi \\
&= e^{\lambda}h'(y-1)(1+O(1/y)).
\end{aligned}$$

This equation implies condition (6). □

D. *Appendix. Proof of Lemma 7.1*

Without loss of generality we assume that $a = 0$ and $b = 1$. For any $\alpha < \beta$ we define the set

$$U_{\alpha,\beta} = \{x \in [0,1] : \alpha < g(x) < \beta\}.$$

The lemma will be proven if for any $\alpha < \beta$ we show that the inequality $|U_{\alpha,\beta}| > 0$ (where $|\cdot|$ is the Lebesgue measure) implies the estimate $|U_{\alpha,\beta}| > 1/2$. Indeed, if g does not coincide almost everywhere with a constant, there exist $\alpha < \beta < \alpha' < \beta'$ such that $|U_{\alpha,\beta}| > 0$ and $|U_{\alpha',\beta'}| > 0$. Then we obtain

$$|U_{\alpha,\beta}| + |U_{\alpha',\beta'}| > 1 = |[0,1]|.$$

Below we fix α and β and, for brevity, use the notation U instead of $U_{\alpha,\beta}$.

Suppose that $|U| > 0$. Then there exists a density point $x_0 \in (0,1)$. This means that for any $\varepsilon > 0$ there exists an interval $I_\varepsilon \subset [0,1]$, $x_0 \in I_\varepsilon$ such that

$$\frac{|U \cap I_\varepsilon|}{|I_\varepsilon|} > 1 - \varepsilon.$$

We put $h_0(1) = A$ and $h_1(0) = B$. Since $\lambda_0 + \lambda_1 > 1$, we have $A > B$. Now we consider the sequence of intervals

$$K_0, K_1, \ldots, K_N \subset [0,1],$$

constructed inductively as follows.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press

(0) $K_0 = I_\varepsilon$.
(1) If $K_n \subset (0,\, A)$, we put $K_{n+1} = h_0^{-1}(K_n)$.
(2) If $K_n \not\subset (0,\, A)$ but $K_n \subset (B,\, 1)$, we put $K_{n+1} = h_1^{-1}(K_n)$.
(3) We finish if $K_n \not\subset (0,\, A)$ and $K_n \not\subset (B,\, 1)$ (i.e., $(B,\, A) \subset K_n$).

Since $\lambda_0,\, \lambda_1 < 1$, the sequence $|K_0|, |K_1|, \ldots$ grows exponentially. Therefore the sequence stops at some finite step $n = N$. Note also that because of the condition $g = g \circ h_0 = g \circ h_1$ we have

$$\frac{|U \cap K_n|}{|K_n|} > 1 - \varepsilon \quad \text{for all } n = 0, \ldots, N.$$

The estimate $|K_N \backslash U| < \varepsilon |K_N| \le \varepsilon$ and the inclusion $(B,\, A) \subset K_N$ imply that $|(B,\, A) \backslash U| \le \varepsilon$. Therefore

$$\frac{|U \cap (B,\, A)|}{|(B,\, A)|} > 1 - \frac{\varepsilon}{A - B}.$$

At least one of the inequalities $|(0,\, A)| > 1/2$, $|(B,\, 1)| > 1/2$ holds. For definiteness let it be the second one. Applying the map h_0^{-1}, we have

$$\frac{|U \cap (C,\, 1)|}{|(C,\, 1)|} > 1 - \frac{\varepsilon}{A - B}, \quad C = h_0^{-1}(B).$$

Now we consider the interval

$$J = h_1^{-s}((C,\, 1)) = (h_1^{-s}(C),\, 1)$$

with minimal $s \in \{0,\, 1,\, 2,\, \ldots\}$ such that $B \in J$. Then

$$|J| > |(B,\, 1)| > 1/2. \tag{D.1}$$

The estimate $|U \cap J|/|J| > 1 - \varepsilon/(A - B)$ implies that

$$|U| \ge |U \cap J| > |J| \left(1 - \frac{\varepsilon}{A - B}\right).$$

By (D.1) the last expression is greater than $1/2$ for sufficiently small ε. Hence we obtain the required inequality $|U| > 1/2$. □

References

[1] V. I. Arnold, V. V. Kozlov and A. I. Neishtadt. *Mathematical Aspects of Classical and Celestial Mechanics*. Springer, Berlin, 2006.

[2] J. R. Cary, D. F. Escande and J. L. Tennyson. Adiabatic invariant change due to separatrix crossing. *Phys. Rev. A* **34**(5) (1986), 4256–4275.

[3] U. Krengel. *Ergodic Theorems*. Walter de Gruyter, Berlin, 1985.

[4] A. I. Neishtadt. Passage through a separatrix in a resonance problem with a slowly-varying parameter. *Prikl. Mat. Mekh.* **39**(4) (1975), 621–632 (Engl. transl. *J. Appl. Math. Mech.* **39**(4) (1975), 594–605).

[5] A. I. Neishtadt. Change of an adiabatic invariant at a separatrix. *Fiz. Plazmy* **12**(8) (1986), 992–1000 (Engl. transl. *Sov. J. Plasma Phys.* **12** (1986), 568–573).

[6] A. V. Timofeev. On the constancy of an adiabatic invariant when the nature of the motion changes. *Zh. Ehksper. Teor. Fiz.* **75**(4) (1978), 1303–1308 (Engl. transl. *Sov. Phys., JETP* **48** (1978), 656–659).

[7] A. M. Vershik. Multivalued mappings with invariant measure (polymorphisms) and Markov operators. *Zap. Nauchn. Semin. LOMI* **72** (1977), 26–61 (Engl. transl. *J. Sov. Math.* **23** (1983), 2243–2266).

[8] A. M. Vershik. Polymorphisms, Markov processes, and quasi-similarity. *Discrete Contin. Dyn. Syst.* **13**(5) (2005), 1305–1324.

https://doi.org/10.1017/S0143385709001060 Published online by Cambridge University Press