Acta Math., 194 (2005), 37–132

Physical measures for partially hyperbolic surface endomorphisms

by

MASATO TSUJI

Hokkaido University
Sapporo, Japan

1. Introduction

In the study of smooth dynamical systems from the standpoint of ergodic theory, one of the most fundamental questions is whether the following preferable picture is true for almost all of them: The asymptotic distribution of the orbit for Lebesgue almost every initial point exists and coincides with one of the finitely many ergodic invariant measures that are given for the dynamical system. The answer is expected to be affirmative in general [14]. However, it seems far beyond the scope of present research to answer the question in the general setting. The purpose of this paper is to provide an affirmative answer to the question in the case of partially hyperbolic endomorphisms on surfaces with one-dimensional unstable subbundle.

Let M be the two-dimensional torus $\mathbf{T}=\mathbf{R}^2/\mathbf{Z}^2$ or, more generally, a region on the torus $\mathbf{T}$ whose boundary consists of finitely many simple closed C^2-curves: e.g. an annulus $(\mathbf{R}/\mathbf{Z})\times\left[-\frac{1}{3},\frac{1}{3}\right]$. We equip M with the Riemannian metric $\|\cdot\|$ and the Lebesgue measure $\mathbf{m}$ that are induced by the standard ones on the Euclidean space $\mathbf{R}^2$ in an obvious manner. We call a C^1-mapping $F\colon M\to M$ a *partially hyperbolic endomorphism* if there are positive constants λ and c and a continuous decomposition of the tangent bundle $TM=\mathbf{E}^c\oplus\mathbf{E}^u$ with $\dim\mathbf{E}^c=\dim\mathbf{E}^u=1$ such that

(i) $\|DF^n|_{\mathbf{E}^u(z)}\|>\exp(\lambda n-c)$;

(ii) $\|DF^n|_{\mathbf{E}^c(z)}\|<\exp(-\lambda n+c)\|DF^n|_{\mathbf{E}^u(z)}\|$

for all $z\in M$ and $n\geqslant 0$. The subbundles $\mathbf{E}^c$ and $\mathbf{E}^u$ are called the central and unstable subbundle, respectively. Notice that we do not require these subbundles to be invariant in the definition, though the central subbundle $\mathbf{E}^c$ turns out to be forward invariant from the condition (ii). The totality of partially hyperbolic C^r-endomorphisms on M is an open subset in the space $C^r(M,M)$, provided $r\geqslant 1$.

An invariant Borel probability measure μ for a dynamical system $F: M \to M$ is said to be a *physical measure* if its basin of attraction,

$$\mathcal{B}(\mu) = \mathcal{B}(\mu; F) := \left\{ z \in M \;\middle|\; \frac{1}{n}\sum_{i=0}^{n-1} \delta_{F^i(z)} \to \mu \text{ weakly as } n \to \infty \right\},$$

has positive Lebesgue measure. One of the main results of this paper is the following theorem:

THEOREM 1.1. *A partially hyperbolic C^r-endomorphism on M generically admits finitely many ergodic physical measures whose union of basins of attraction has total Lebesgue measure, provided that $r \geqslant 19$.*

More detailed versions of this theorem will be given in the next section. Here we intend to explain the new idea behind the results of this paper. The readers should notice that we do *not* (and will *not*) claim that the physical measures in the theorem above are hyperbolic. Instead, we will show that the physical measures for generic partially hyperbolic endomorphisms have nice properties even if they are not hyperbolic. This is the novelty of the argument in this paper.

Let us consider a partially hyperbolic endomorphism F on M. The Lyapunov exponent of F takes two distinct values at each point: The larger is positive and the smaller indefinite. The latter is called the central Lyapunov exponent, as it is attained by the vectors in the central subbundle. An invariant measure for F is *hyperbolic* if the central Lyapunov exponent is non-zero at almost every point with respect to it. In the former part of this paper, we study hyperbolic invariant measures for partially hyperbolic endomorphisms using the techniques in the Pesin theory or the smooth ergodic theory. And, as the conclusion, we show that the following hold under some generic conditions on F: *For any $\chi > 0$, there are only finitely many ergodic physical measures whose central Lyapunov exponents are larger than χ in absolute value. Further, if the complement X of the union of the basins of attraction of such physical measures has positive Lebesgue measure, and if a measure μ is a weak limit point of the sequence*

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbf{m}|_X \circ F^{-i}, \quad n = 1, 2, \ldots, \tag{1}$$

($\mathbf{m}|_X$ is the restriction of $\mathbf{m}$ to X), then the absolute value of the central Lyapunov exponent is not larger than χ at almost every point with respect to μ. Though these facts are far from trivial, the argument in the proof does not deviate far from the existing ones in the smooth ergodic theory.

The key claim in our argument is the following: *If the number χ is small enough and if F satisfies some additional generic conditions, then a measure μ as in the preceding*

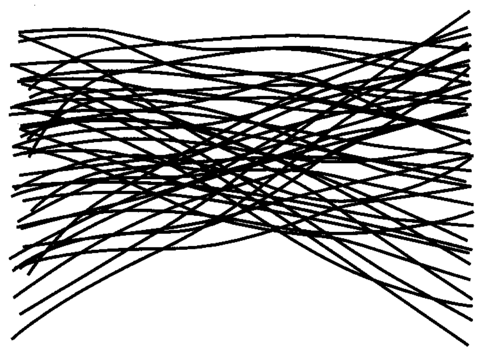

Fig. 1. The curve $F^n(\gamma)$.

paragraph is absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$. *Further, the density* $d\mu/d\mathbf{m}$ *satisfies some regularity conditions (from which we can conclude Theorem* 1.1). This claim might appear unusual, since the measure μ may have neutral or even negative central Lyapunov exponent, while we usually meet absolutely continuous invariant measures as a consequence of the expanding property of dynamical systems in all directions. We can explain it intuitively as follows: As a consequence of the dominating expansion in the unstable directions $\mathbf{E}^u$, the measure μ should have some smoothness or uniformity in those directions. In fact, we can show that the natural extensions of μ and its ergodic components to the inverse limit are absolutely continuous along the (one-dimensional) unstable manifolds. So, for each ergodic component μ' of μ, we can cut a curve γ out of an unstable manifold so that μ' is attained as a weak limit point of the sequence $n^{-1}\sum_{i=0}^{n-1}\nu_\gamma\circ F^{-i}$, $n=1,2,\ldots$, where ν_γ is a smooth measure on γ. Since F expands the curve γ uniformly, the image $F^n(\gamma)$ for large n should be a very long curve which is transversal to the central subbundle $\mathbf{E}^c$. Imagine looking into a small neighborhood of a point in the support of μ'. The image $F^n(\gamma)$ should appear as a bunch of short pieces of curve in that neighborhood; see Figure 1.

The number of the pieces of curve should grow exponentially as n gets large. And they would not concentrate strongly in the central direction, as the central Lyapunov exponent is nearly neutral almost everywhere with respect to μ'. These consequences suggest that the ergodic component μ' should have some smoothness or uniformity in the central direction as well as in the unstable direction, and so the measure μ should be absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$.

On the technical side, an important idea in the proof of the key claim is that we look at the angles between the short pieces of curve mentioned above rather than their positions. As we perturb the mapping F, it turns out that we can control the angles between those pieces of curve to some extent, though we cannot control their positions by the usual problem of interference. And we can show that the pieces of curve satisfy some transversality condition generically. In order to show the conclusion of the key claim, we relate that transversality condition to absolute continuity of the measure μ. To this end, we make use of an idea in the paper [15] by Peres and Solomyak with some modification. We will illustrate the idea in the beginning of §6 by using a simple example. Actually we have used the same idea in our previous paper [24], which can be regarded as a study for this work. Lastly, the author would like to note that the idea in [15] can be traced back to the papers of Falconer [5] and Simon and Pollicot [17].

Acknowledgment. I would like to thank Jérôme Buzzi, Viviane Baladi, Artur Avila and Mitsuhiro Shishikura for valuable comments in writing this paper.

2. Statement of the main results

Let $\mathcal{PH}^r$ be the set of partially hyperbolic C^r-endomorphisms on M, and $\mathcal{PH}_0^r$ the subset of those without critical points. The subset $\mathcal{R}^r \subset \mathcal{PH}^r$ is the totality of mappings $F \in \mathcal{PH}^r$ that satisfy the following two conditions:

(a) F admits a finite collection of ergodic physical measures whose union of basins of attraction has total Lebesgue measure on M;

(b) A physical measure for F is absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$ if the sum of its Lyapunov exponents is positive.

In this paper, we claim that almost all partially hyperbolic C^r-endomorphisms on M satisfy the conditions (a) and (b) above, or, in other words, belong to the subset $\mathcal{R}^r$. The former part of our main result is stated as follows:

THEOREM 2.1. (I) *The subset $\mathcal{R}^r$ is a residual subset in $\mathcal{PH}^r$, provided $r \geqslant 19$.*

(II) *The intersection $\mathcal{R}^r \cap \mathcal{PH}_0^r$ is a residual subset in $\mathcal{PH}_0^r$, provided $r \geqslant 2$.*

The conclusions of this theorem mean that the complement of the subset $\mathcal{R}^r$ is a meager subset in the sense of Baire's category argument. However, the recent progress in dynamical system theory has thrown serious doubt that the notion of genericity based on Baire's category argument may not have its literal meaning. In fact, it can happen that the dynamical systems in some meager subset appear as subsets with positive Lebesgue measure in the parameter spaces of typical families. For example, compare Jakobson's theorem [23] and the density of Axiom A [12], [19] in one-dimensional dynamical systems. For this reason, we dare to state our claim also in a measure-theoretical framework, though no measure-theoretical definition that corresponds to the notion of genericity has been firmly established yet.

Let B be a Banach space. Let $\tau_v : B \to B$ be the translation by $v \in B$, that is, $\tau_v(x) = x + v$. A Borel finite measure $\mathcal{M}$ on B is said to be *quasi-invariant* along a linear subspace

$L \subset B$ if $\mathcal{M}\circ\tau_v^{-1}$ is equivalent to $\mathcal{M}$ for any $v \in L$. In the case where B is finite-dimensional, a Borel finite measure on B is equivalent to the Lebesgue measure if and only if it is quasi-invariant along the whole space B. But, unfortunately enough, it is known that no Borel finite measures on an infinite-dimensional Banach space are quasi-invariant along the whole space [6]. This is one of the reasons why we do not have obvious definitions for concepts like *Lebesgue almost everywhere* in the cases of infinite-dimensional Banach spaces or Banach manifolds such as the space $C^r(M, M)$. However, there may be Borel finite measures on B that are quasi-invariant along dense subspaces. In fact, on the Banach space $C^r(M, \mathbf{R}^2)$, there exist Borel finite measures that are quasi-invariant along the dense subspace $C^{r+2}(M, \mathbf{R}^2)$ (Lemma 3.18). For integers $s \geqslant r \geqslant 1$, let $\mathcal{Q}_s^r$ be the set of Borel probability measures on $C^r(M, \mathbf{R}^2)$ that are quasi-invariant along the subspace $C^s(M, \mathbf{R}^2)$ and regard the measures in these sets as substitutions for the (non-existing) Lebesgue measure.

Let us consider the space $C^r(M, \mathbf{T})$ of C^r-mappings from M to the torus $\mathbf{T}$, which contains the space $C^r(M, M)$ of C^r-endomorphisms on M. For a mapping G in $C^r(M, \mathbf{T})$, we consider the mapping

$$\begin{aligned} \Phi_G : C^r(M, \mathbf{R}^2) &\longrightarrow C^r(M, \mathbf{T}), \\ F &\longmapsto G + F. \end{aligned} \tag{2}$$

We now introduce the following notions:

Definition. A subset $\mathcal{X} \subset C^r(M, M)$ is *shy* with respect to a measure $\mathcal{M}$ on $C^r(M, \mathbf{R}^2)$ if $\Phi_G^{-1}(\mathcal{X})$ is a null subset with respect to $\mathcal{M}$ for any $G \in C^r(M, \mathbf{T})$.

Definition. A subset $\mathcal{X} \subset C^r(M, M)$ is *timid* for the class $\mathcal{Q}_s^r$ of measures if $\mathcal{Q}_s^r$ is non-empty and if $\mathcal{X}$ is shy with respect to *all* measures in $\mathcal{Q}_s^r$.

Remark. The former of the definitions above is a slight modification of that of shyness introduced by Hunt, Sauer and Yorke [9], [10]. By the definitions, a subset $\mathcal{X} \subset C^r(M, M)$ is shy with respect to some compactly supported measure $\mathcal{M}$ in the sense above if and only if the inverse image $\Phi_G^{-1}(\mathcal{X})$ for some (and thus any) $G \in C^r(M, \mathbf{T})$ is shy in the sense of Hunt, Sauer and Yorke. Note that a controversy to the notion of shyness is given in the paper [21] of Stinchcombe.

Put $\mathcal{S}^r := \mathcal{PH}^r \setminus \mathcal{R}^r$. The latter part of our main result is stated as follows:

THEOREM 2.2. (I) *The subset $\mathcal{S}^r$ is shy with respect to a measure $\mathcal{M}_s$ in $\mathcal{Q}_{s-1}^r$ if the integers $r \geqslant 2$ and $s \geqslant r+3$ satisfy*

$$(r-2)(r+1) < (r-\nu-2)\left(2(r-3) - \frac{2s-r-\nu+1}{\nu}\right) \tag{3}$$

for some integer $3 \leqslant \nu \leqslant r-2$. *Moreover,* $\mathcal{S}^r$ *is timid for* $\mathcal{Q}^r_{s-1}$ *if* $r \geqslant 2$ *and* $s \geqslant r+3$ *satisfy the condition* (3) *with* s *replaced by* $s+2$ *for some integer* $3 \leqslant \nu \leqslant r-2$.

(II) $\mathcal{S}^r \cap \mathcal{PH}^r_0$ *is timid for* $\mathcal{Q}^r_s$ *for any* $r \geqslant 2$ *and* $s \geqslant r+2$.

Remark. The measure $\mathcal{M}_s$ in the claim (I) above will be constructed explicitly as a Gaussian measure.

Remark. The inequality (3) holds for the combinations $(r, s, \nu) = (19, 22, 3)$ and $(21, 26, 3)$, for example. But it does not hold for any $s \geqslant r+3$ and $3 \leqslant \nu \leqslant r-2$ unless $r \geqslant 19$.

As an advantage of the measure-theoretical notion of timidity introduced above, we can derive the following corollary on the families of mappings in $\mathcal{PH}^r$, whose proof is given in the appendix. Let us regard the space $C^r(M \times [-1,1]^k, M)$ as that of C^r-families of endomorphisms on M with parameter space $[-1,1]^k$. We can introduce the notion of shyness and timidity for the Borel subsets in this space in the same way as we did for those in $C^r(M, M)$. Let $\mathbf{m}_{\mathbf{R}^k}$ be the Lebesgue measure on $\mathbf{R}^k$.

COROLLARY 2.3. *The set of* C^r*-families* $F(z, \mathbf{t})$ *in* $C^r(M \times [-1,1]^k, M)$ *satisfying*

$$\mathbf{m}_{\mathbf{R}^k}(\{\mathbf{t} \in [-1,1]^k \mid F(\,\cdot\,, \mathbf{t}) \in \mathcal{S}^r\}) > 0$$

is timid with respect to the class of Borel finite measures on $C^r(M \times [-1,1]^k, \mathbf{R}^2)$ *that are quasi-invariant along the subspace* $C^{s-1}(M \times [-1,1]^k, \mathbf{R}^2)$*, provided that the integers* $r \geqslant 2$ *and* $s \geqslant r+3$ *satisfy the condition* (3) *with* s *replaced by* $s+2$ *for some integer* $3 \leqslant \nu \leqslant r-2$.

We give a few comments on the main results above. The restriction that the surface M is a region on the torus is actually not very essential. We could prove Theorem 2.1 with M being a general compact surface by modifying the proof slightly. The main reason for this restriction is the difficulty in generalizing the notion of shyness and timidity to the spaces of endomorphisms on general compact surfaces. Since the definitions depend heavily on the linear structure of the space $C^r(M, \mathbf{R}^2)$, we hardly know how to modify these notions naturally so that it is consistent under the non-linear coordinate transformations. The generalization or modification of these notions should be an important issue in the future. Besides, the restriction on M simplifies the proof considerably and does not exclude the interesting examples such as the so-called Viana–Alves maps [1], [25].

The assumptions on differentiability in the main theorems are crucial in our argument, especially in the part where we consider the influence of the critical points on the dynamics. We do not know whether they are technical ones or not.

As we called attention to in the introduction, the main theorems tell nothing about hyperbolicity of the physical measures. Of course, it is natural to expect that the physical measures are hyperbolic generically. The author thinks it not too optimistic to expect that $\mathcal{R}^r$ contains an open dense subset of $\mathcal{PH}^r$ in which the physical measures for the mappings are hyperbolic and depend on the mapping continuously.

Generalization of the main theorems to partially hyperbolic diffeomorphisms on higher-dimensional manifolds is an interesting subject to study. Our argument on physical measures with nearly neutral central Lyapunov exponent seems to be complementary to the recent works [2] and [3] of Alves, Bonatti and Viana. However, as far as the author understands, there exist essential difficulties in the case where the dimension of the central subbundle is higher than one.

The plan of this paper is as follows: We give some preliminaries in §3. We first define some basic notation and then introduce the notions of *admissible curve* and *admissible measure*, which play central roles in our argument. The former is taken from the paper [25] of Viana with slight modification and the latter is a corresponding notion for measures. Next we introduce two conditions on partially hyperbolic endomorphisms, namely, *the transversality condition on unstable cones* and *the no flat contact condition.* At the end of §3, we shall give a concrete plan of the proof of the main theorems using the terminology introduced in this section. In §4, we study hyperbolic physical measures using the Pesin theory. §5 is devoted to basic estimates on the distortion of the iterates of partially hyperbolic endomorphisms. Then we go into the main part of this paper, which consists of three mutually independent sections. In §6, we prove that a partially hyperbolic endomorphism belongs to the subset $\mathcal{R}^r$ if it satisfies the two conditions above. In §§ 7 and 8 respectively, we prove that each of the two conditions holds for almost all partially hyperbolic endomorphisms.

3. Preliminaries

In this section, we prepare some notation, definitions and basic lemmas that we shall use frequently in the following sections.

3.1. Notation

For a tangent vector $v \in TM$, $v^{\perp}$ denotes the tangent vector that is obtained by rotating v by the right angle in the counter-clockwise direction. For two tangent vectors u and v, $\angle(u, v)$ denotes the angle between them even if they belong to the tangent spaces at different points. Let $\exp_z : T_z\mathbf{T} \to \mathbf{T}$ be the exponential mapping, which is defined simply by $\exp_z(v)=z+v$ in our case. For a point z in the torus $\mathbf{T}$ (or in some metric space, more generally) and a positive number δ, let $\mathbf{B}(z,\delta)$ be the open disk with center at z and radius δ. Likewise, for a subset X, let $\mathbf{B}(X,\delta)$ be its open δ-neighborhood. For a positive number δ, we define a lattice $\mathbf{L}(\delta)$ as the subset of points (x,y) in $\mathbf{T}$ whose components, x and y, are multiples of $1/([1/\delta]+1)$, so that the disks $\mathbf{B}(z,\delta)$ for points $z\in\mathbf{L}(\delta)$ cover the torus $\mathbf{T}$.

Throughout this paper, we assume $r\geqslant 2$. Let $F\colon M\to M$ be a C^r-mapping. The set of critical points of F is denoted by $\mathcal{C}(F)$. For a tangent vector $v\in T_zM$ at a point $z\in M$, we define

$$D_*F(z,v)=\frac{\|DF_z(v)\|}{\|v\|}\quad \text{if } v\neq\mathbf{0},$$

and

$$D^*F(z,v)=\frac{\det DF_z}{D_*F(z,v)}\quad \text{if } DF(v)\neq\mathbf{0}.$$

Remark. If $v\neq\mathbf{0}$ and $DF(v)\neq\mathbf{0}$, we can take orthonormal bases $(v/\|v\|, v^\perp/\|v^\perp\|)$ on T_zM and $(DF(v)/\|DF(v)\|, DF(v)^\perp/\|DF(v)^\perp\|)$ on $T_{F(z)}M$. Then the representation matrix of $DF_z\colon T_zM\to T_{f(z)}M$ with respect to these bases is an upper triangular matrix with $D_*F(z,v)$ and $D^*F(z,v)$ on the diagonal.

Note that we have $|D^*F(z,v)|=\|(DF)^*(v^*)\|/\|v^*\|$ for any cotangent vector $v^*\neq\mathbf{0}$ at $F(z)$ that is normal to $DF(v)$. We shall write $D_*F(v)$ and $D^*F(v)$ for $D_*F(z,v)$ and $D^*F(z,v)$, respectively, in places where the base point z is clear from the context.

For a C^r-mapping $F\colon M\to\mathbf{R}^2$, the C^r-norm of F is defined by

$$\|F\|_{C^r}=\max_{z\in M}\max_{0\leqslant a+b\leqslant r}\left\|\frac{\partial^{a+b}F}{\partial^a x\,\partial^b y}(z)\right\|,$$

where (x,y) is the coordinate on $\mathbf{T}$ that is induced by the standard one on $\mathbf{R}^2$. Similarly, for C^r-mappings F and G in $C^r(M,\mathbf{T})$, the C^r-distance is defined by

$$d_{C^r}(F,G)=\max_{z\in M}\max\left\{d(F(z),G(z)),\ \max_{1\leqslant a+b\leqslant r}\left\|\frac{\partial^{a+b}F}{\partial^a x\,\partial^b y}(z)-\frac{\partial^{a+b}G}{\partial^a x\,\partial^b y}(z)\right\|\right\}.$$

3.2. Some open subsets in $\mathcal{PH}^r$

In this subsection, we introduce some bounded open subsets in $\mathcal{PH}^r$ whose elements enjoy certain estimates uniformly. As we will see, we can restrict ourselves to such open subsets in proving the main theorems. This simplifies the argument considerably.

Let $\mathcal{S}_0^r$ be the subset of mappings F in $\mathcal{PH}^r$ that *violate* either of the conditions:

(A1) The image $F(M)$ is contained in the interior of M;

(A2) The function $z \mapsto \det DF_z$ has 0 as its regular value;

(A3) The restriction of F to the critical set $\mathcal{C}(F)$ is transversal to $\mathcal{C}(F)$.

Notice that the conditions (A2) and (A3) are trivial if the mapping F has no critical points. To prove the following lemma, we have only to apply Thom's jet transversality theorem [7] and its measure-theoretical version [22, Theorem C].

LEMMA 3.1. *The subset $\mathcal{S}_0^r$ is a closed nowhere dense subset in $\mathcal{PH}^r$ and shy with respect to any measure in $\mathcal{Q}_s^r$ for $s \geqslant r \geqslant 2$.*

Remark. The terminology in [22] is different from that in this paper. But we can put Theorem C and other results in [22] into our terminology without difficulty.

Consider a C^r-mapping $F_\sharp$ in $\mathcal{PH}^r$ and let $TM = \mathbf{E}^c \oplus \mathbf{E}^u$ be a decomposition of the tangent bundle which satisfies the conditions in the definition that $F = F_\sharp$ is a partially hyperbolic endomorphism. Notice that, although the central subbundle $\mathbf{E}^c$ is uniquely determined by the conditions in the definition, the unstable subbundle $\mathbf{E}^u$ is *not*. Indeed any continuous subbundle transversal to $\mathbf{E}^c$ satisfies the conditions in the definition, possibly with different constants λ and c. Making use of this arbitrariness, we can assume that $\mathbf{E}^u$ is a C^∞-subbundle. Further, by taking $\mathbf{E}^u$ nearly orthogonal to $\mathbf{E}^c$ and by changing the constants λ and c, we can assume that there exist positive-valued C^∞-functions θ^c and θ^u on M such that the cone fields

$$\mathbf{S}^u(z) = \{v \in T_z M \setminus \{0\} \mid \angle(v, \mathbf{E}^u(z)) \leqslant \theta^u(z)\},$$
$$\mathbf{S}^c(z) = \{v \in T_z M \setminus \{0\} \mid \angle(v^\perp, \mathbf{E}^u(z)) \leqslant \theta^c(z)\}$$

satisfy the following conditions at every point $z \in M$:

(B1) $\mathbf{S}^c(z) \cap \mathbf{S}^u(z) = \varnothing$;

(B2) $\mathbf{E}^c(z) \setminus \{0\}$ is contained in the interior of the cone $\mathbf{S}^c(z)$;

(B3) $DF_\sharp(\mathbf{S}^u(z))$ is contained in the interior of $\mathbf{S}^u(F_\sharp(z))$;

(B4) $(DF_\sharp)_z^{-1}(\mathbf{S}^c(F_\sharp(z)))$ is contained in the interior of $\mathbf{S}^c(z)$;

(B5) For any $v \in \mathbf{S}^u(z)$ and $n \geqslant 1$, we have

(i) $\|D_* F_\sharp^n(z, v)\| > \exp(\lambda n - c)$;

(ii) $\|D^* F_\sharp^n(z, v)\| < \exp(-\lambda n + c) \|D_* F_\sharp^n(z, v)\|$.

Suppose that the mapping $F_\sharp$ does *not* belong to $\mathcal{S}_0^r$. Then we can take positive constants λ and c, a small number $\varrho > 0$ and a large number $\Lambda > c$ such that the following conditions hold for any C^r-mapping F satisfying $d_{C^r}(F, F_\sharp) < 2\varrho$:

(C1) The conditions (B3), (B4) and (B5) with $F_\sharp$ replaced by F hold;

(C2) The parallel translation of $\mathbf{E}^c(F_\natural(z))$ to $F(z)$ is contained in $\mathbf{S}^c(F(z))\cup\{0\}$ for any $z\in M$;

(C3) $d(F(M),\partial M)>\varrho$;

(C4) The function $z\mapsto\det DF_z$ has no critical points on $\mathbf{B}(\mathcal{C}(F),\varrho)$, and it holds that $|\det DF_z|>\varrho d(z,\mathcal{C}(F))$ for $z\in\mathbf{B}(\mathcal{C}(F),\varrho)$;

(C5) If a point $z\in M$ satisfies $d(z,w_1)<\varrho$ and $d(F(z),w_2)<\varrho$ for some points $w_1,w_2\in\mathcal{C}(F)$ and if $v\in\mathbf{S}^u(z)$, then the angle between $DF(v)$ and the tangent vector of $\mathcal{C}(F)$ at w_2 is larger than ϱ;

(C6) $\|DF_z\|<\Lambda$ for any $z\in M$.

We can choose countably many pairs of a C^r-mapping $F_\natural$ in $\mathcal{PH}^r\setminus\mathcal{S}_0^r$ and a positive number ϱ as above so that the corresponding open subsets

$$\mathcal{U}=\{F\in C^r(M,M)\mid d_{C^r}(F_\natural,F)<\varrho\}$$

cover $\mathcal{PH}^r\setminus\mathcal{S}_0^r$. In order to prove the main theorems, Theorems 2.1 and 2.2, it is enough to prove them by restricting ourselves to an arbitrary such open subset $\mathcal{U}$. For this reason, we henceforth fix a C^r-mapping $F_\natural$ in $\mathcal{PH}^r\setminus\mathcal{S}_0$, subbundles $\mathbf{E}^c$ and $\mathbf{E}^u$, C^∞-functions θ^c and θ^u, cone fields $\mathbf{S}^c(\cdot)$ and $\mathbf{S}^u(\cdot)$, and positive numbers ϱ, Λ, λ and c as above, and consider the mappings in the corresponding open subset $\mathcal{U}$.

3.3. Remarks on the notation for constants

In this paper, we shall introduce various constants that depend only on

(1) the objects that we have just fixed at the end of the last subsection;

(2) the integer $r\geqslant 2$.

In order to distinguish such constants, we make it a rule to write them by symbols with subscript g. Obeying this rule, we shall write λ_g, c_g, ϱ_g and Λ_g for the constants λ, c, ϱ and Λ hereafter (and we will use the symbols λ, c, ϱ and Λ for other purposes). Notice that, once a constant is denoted by a symbol with subscript g, we mean that it is a constant of this kind. In order to save symbols for constants, we shall frequently use a generic symbol C_g for large positive constants of this kind. Note that the value of the constants denoted by C_g may be different from place to place even in a single expression. For instance, ridiculous expressions like $2C_g<C_g$ can be true, though we shall not really meet such ones. Also note that we shall omit the phrases on the choice of the constants denoted by C_g in most cases.

For example, we can take a constant $A_g>0$ such that

$$A_g^{-1}\frac{|D^*F^n(z,w)|}{D_*F^n(z,w)}\leqslant\frac{\angle(DF^n(u),DF^n(v))}{\angle(u,v)}\leqslant A_g\frac{|D^*F^n(z,w)|}{D_*F^n(z,w)} \tag{4}$$

for any $z\in M$, $n\geqslant 1$ and $u, v, w\notin \mathbf{S}^c(z)\cup\{\mathbf{0}\}$. We shall use the following relations frequently: For any $F\in\mathcal{U}$, $z\in M$, $v\in\mathbf{S}^u(z)$ and $n\geqslant 1$, we have

$$C_g^{-1}d(z,\mathcal{C}(F))\leqslant|\det DF_z|\leqslant\exp(\Lambda_g)\|D^*F(z,v)\|\leqslant C_g d(z,\mathcal{C}(F)), \tag{5}$$

$$C_g^{-1}<\|D_*F^n(z,v)\|\leqslant\|DF_z^n\|\leqslant C_g\|D_*F^n(z,v)\| \tag{6}$$

and, if $z\notin\mathcal{C}(F)$, also

$$C_g^{-1}\|D^*F^n(z,v)\|\leqslant\|(DF_z^n)^{-1}\|^{-1}\leqslant\|D^*F^n(z,v)\|. \tag{7}$$

3.4. Admissible curves

In this subsection, we introduce the notion of admissible curve. From the forward invariance of the unstable cones $\mathbf{S}^u$ or the condition (B3) with $F_\sharp$ replaced by F, the mappings in the set $\mathcal{U}$ preserve the class of C^1-curves whose tangent vectors belong to $\mathbf{S}^u$. We shall investigate such a class of curves and find a subclass which is uniformly bounded in C^{r-1}-sense and essentially invariant under the iterates of mappings in $\mathcal{U}$. We shall call the curves in this subclass admissible curves.

In this paper, we always assume that the curves are regular and parameterized by length. Let $\gamma\colon[0,a]\to M$ be a C^r-curve such that $\gamma'(t)\in\mathbf{S}^u(\gamma(t))$ for $t\in[0,a]$. As we assume $\|\gamma'(t)\|\equiv 1$, the second differential of γ is written in the form

$$\frac{d^2}{dt^2}\gamma(t)=d^2\gamma(t)(\gamma'(t))^\perp,$$

where $d^2\gamma\colon[0,a]\to\mathbf{R}$ is a C^{r-2}-function. We define $d^k\gamma(t)$ for $3\leqslant k\leqslant r$ as the $(k-2)$nd differential of the function $d^2\gamma(t)$. Let $d^1\gamma(t)$ be the differential $\gamma'(t)$, for convenience.

Let $F_*\gamma\colon[0,a']\to M$ be the image of the curve γ under a mapping $F\in\mathcal{U}$. Notice that $F_*\gamma$ is not simply the composition $F\circ\gamma$, because we assume $F_*\gamma$ to be parameterized by length. The right relation between γ and $F_*\gamma$ is given by

$$F_*\gamma(p(t))=F(\gamma(t)), \tag{8}$$

where $p\colon[0,a]\to[0,a']$ is the unique C^r-diffeomorphism satisfying $p(0)=0$ and $dp(t)/dt=D_*F(\gamma(t),\gamma'(t))$. Differentiating both sides of (8), we get the formula

$$D_*F(\gamma(t),\gamma'(t))\cdot(F_*\gamma)'(p(t))=DF_{\gamma(t)}(\gamma'(t))$$

for $t\in[0,a]$. Differentiating both sides again and considering the components normal to $(F_*\gamma)'(p(t))$, we get

$$d^2F_*\gamma(p(t))=\frac{D^*F(\gamma(t),\gamma'(t))}{D_*F(\gamma(t),\gamma'(t))^2}\,d^2\gamma(t)+\frac{Q_2(\gamma(t),\gamma'(t);F)}{D_*F(\gamma(t),\gamma'(t))^3}, \tag{9}$$

where

$$Q_2(a,b;F)=(D^2F_a(b,b),(DF_a(b))^{\perp}).$$

Note that $Q_2(a,b;F)$ is a polynomial of the components of the unit vector b whose coefficients are polynomials of the differentials of F at a up to the second order. Likewise, examining the differentials of both sides of (9) by using the relation

$$\frac{d}{dt}D_*F(\gamma(t),\gamma'(t))=\frac{\frac{d}{dt}\|DF_{\gamma(t)}(\gamma'(t))\|^2}{2D_*F(\gamma(t),\gamma'(t))},$$

we obtain, for $3\leqslant k\leqslant r$,

$$d^kF_*\gamma(p(t))=\frac{D^*F(\gamma(t),\gamma'(t))}{D_*F(\gamma(t),\gamma'(t))^k}d^k\gamma(t)+\frac{Q_k(\gamma(t),\gamma'(t),\{d^i\gamma(t)\}_{i=2}^{k-1};F)}{D_*F(\gamma(t),\gamma'(t))^{3k-3}}, \tag{10}$$

where $Q_k(a,b,\{c_i\}_{i=2}^{k-1};F)$ is a polynomial of the components of the unit vector b and the scalars c_i whose coefficients are polynomials of the differentials of F at a up to the kth order.

Remark. In addition, we can check that $Q_k(a,b,\{c_i\}_{i=2}^{k-1};F)$ for $2\leqslant k\leqslant r$ is written in the form

$$D_*F(a,b)^{2k-3}v^*((D^kF)_a(b,b,...,b))+\widetilde{Q}_k(a,b,\{c_i\}_{i=2}^{k-1};F),$$

where v^* is a unit cotangent vector at the point $F(a)$ that is normal to $DF_a(b)$, and $\widetilde{Q}_k(a,b,\{c_i\}_{i=2}^{k-1};F)$ is a polynomial of the components of b and the scalars c_i whose coefficients are polynomials of the differentials of F at a up to the $(k-1)$st order.

Fix an integer $n_g>0$ such that $n_g\lambda_g-c_g>0$. Then we have the following result:

LEMMA 3.2. *There exist constants $K_g^{(k)}>1$ for $2\leqslant k\leqslant r$ such that, if a curve γ: $[0,a]\to M$ of class C^r satisfies*

(i) *$\gamma'(t)\in\mathbf{S}^u(\gamma(t))$ for $t\in[0,a]$;*

(ii) *$|d^k\gamma(t)|\leqslant K_g^{(k)}$ for $2\leqslant k\leqslant r$ and $t\in[0,a]$,*

then $F_*^n\gamma$ for $n\geqslant n_g$ satisfies the same conditions.

Proof. Consider a C^r-curve γ: $[0,a]\to M$ that satisfies the conditions (i) and (ii), and let $F_*^n\gamma$: $[0,a_n]\to M$ be its image under the iterate F^n. From the formulae (9) and (10), we can see that, for $n_g\leqslant n\leqslant 2n_g$,

$$|d^kF_*^n\gamma(p_n(t))|\leqslant\frac{|D^*F(\gamma(t),\gamma'(t))|}{D_*F(\gamma(t),\gamma'(t))^k}|d^k\gamma(t)|+R(n_g,K_g^{(2)},...,K_g^{(k-1)}), \tag{11}$$

where $p_n: [0,a]\to[0,a_n]$ is the unique diffeomorphism satisfying $p_n(0)=0$ and $dp_n(t)/dt = DF_*^n(\gamma(t),\gamma'(t))$ and where $R(n_g, K_g^{(2)},...,K_g^{(k-1)})$ is a constant that depends only on $n_g, K_g^{(2)},...,K_g^{(k-1)}$ besides the objects that we have already fixed at the end of §3.2. The coefficient of $|d^k\gamma(t)|$ on the right-hand side of the inequality (11) is smaller than $\exp(-n_g\lambda_g+c_g)<1$ from the condition (C1) and the choice of n_g. Thus, if we take a large $K_g^{(k)}$ according to the choice of the constants $K_g^{(2)},...,K_g^{(k-1)}$ in turn for $2\leqslant k\leqslant r$, the conclusion of the lemma holds for $n_g\leqslant n\leqslant 2n_g$. And, employing this repeatedly, we obtain the conclusion for all $n\geqslant n_g$. □

Henceforth we fix the constants $K_g^{(k)}$, $2\leqslant k\leqslant r$, in Lemma 3.2. Now we make the following definition:

Definition. A C^{r-1}-curve $\gamma: [0,a]\to M$ is an *admissible curve* if it satisfies the conditions

(a) $\gamma'(t)\in\mathbf{S}^u(\gamma(t))$ for $t\in[0,a]$;
(b) $|d^k\gamma(t)|\leqslant K_g^{(k)}$ for $2\leqslant k\leqslant r-1$ and $t\in[0,a]$;
(c) the function $d^{r-1}\gamma$ satisfies a Lipschitz condition with the constant $K_g^{(r)}$:

$$|d^{r-1}\gamma(t)-d^{r-1}\gamma(s)|\leqslant K_g^{(r)}|t-s| \quad \text{for any } 0\leqslant s<t\leqslant a.$$

Remark. When $r=2$, the condition (b) above is vacuous and the symbol $|\cdot|$ on the left-hand side of the inequality in the condition (c) should be understood as the norm on $\mathbf{R}^2$. (Recall that we put $d^1\gamma(t)=\gamma'(t)$.)

Note that a C^{r-1}-curve $\gamma: [0,a]\to M$ is an admissible curve if and only if it belongs to the closure, in the space $C^{r-1}([0,a],M)$, of the set of C^r-curves satisfying the conditions (i) and (ii) in Lemma 3.2. Thus we have the following consequence from Lemma 3.2:

COROLLARY 3.3. *If a C^{r-1}-curve γ is admissible, so is $F_*^n\gamma$ for $n\geqslant n_g$.*

For a positive number a, let $\mathcal{A}(a)$ be the set of C^1-curves $\gamma: [0,a]\to M$ of length a such that $\gamma'(t)\in\mathbf{S}^u(\gamma(t))$ for $t\in[0,a]$. For a subset $J\subset(0,\infty)$, we define $\mathcal{A}(J)$ as the disjoint union of $\mathcal{A}(a)$ for $a\in J$:

$$\mathcal{A}(J):=\coprod_{a\in J}\mathcal{A}(a).$$

Also we define

$$\mathbf{A}(J):=\coprod_{a\in J}(\mathcal{A}(a)\times[0,a])\subset\mathcal{A}(J)\times\mathbf{R}.$$

We can regard the space $\mathcal{A}((0,\infty))$ as the totality of C^1-curves whose length are finite and whose tangent vectors are contained in the unstable cone $\mathbf{S}^u$. From the condition

(C1) in the choice of the open neighborhood $\mathcal{U}$, each mapping $F\in\mathcal{U}$ naturally acts on the space $\mathcal{A}((0,\infty))$,

$$F_*: \mathcal{A}((0,\infty)) \longrightarrow \mathcal{A}((0,\infty)),$$
$$\gamma\in\mathcal{A}(a) \longmapsto F_*\gamma\in\mathcal{A}(p(a)),$$

and also on the space $\mathbf{A}((0,\infty))$,

$$F_*: \mathbf{A}((0,\infty)) \longrightarrow \mathbf{A}((0,\infty)),$$
$$(\gamma,t)\in\mathcal{A}(a)\times[0,a] \longmapsto (F_*\gamma,p(t))\in\mathcal{A}(p(a))\times[0,p(a)],$$

where $p: [0,a]\to[0,p(a)]$ is the unique diffeomorphism satisfying $p(0)=0$ and $dp(t)/dt= D_*F(\gamma(t),\gamma'(t))$.

For a positive number a, let $\mathcal{AC}(a)\subset\mathcal{A}(a)$ be the set of admissible curves of length a, and, for a subset $J\subset(0,\infty)$, we put

$$\mathcal{AC}(J) := \coprod_{a\in J}\mathcal{AC}(a)\subset\mathcal{A}(J) \quad \text{and} \quad \mathbf{AC}(J) := \coprod_{a\in J}\mathcal{AC}(a)\times[0,a]\subset\mathbf{A}(J).$$

Note that $\mathcal{AC}(a)$ is a compact subset of $C^{r-1}([0,a],M)$.

We equip the space $\mathcal{AC}((0,\infty))$ with the distance $d_{\mathcal{AC}}$ defined by

$$d_{\mathcal{AC}}(\gamma_1,\gamma_2)=\|\gamma_2-\gamma_1\|_{C^{r-1}}+C|a_2-a_1|$$

for $\gamma_i\in\mathcal{AC}(a_i)$, $i=1,2$, where $\|\gamma_2-\gamma_1\|_{C^{r-1}}$ is

$$\max_{0\leqslant\theta\leqslant\min\{a_1,a_2\}}\max\left\{d(\gamma_2(\theta),\gamma_1(\theta)),\angle(\gamma_1'(\theta),\gamma_2'(\theta)),\max_{2\leqslant k\leqslant r-1}|d^k\gamma_2(\theta)-d^k\gamma_1(\theta)|\right\}$$

and the constant C is defined by

$$C=\max_{2\leqslant k\leqslant r}K_g^{(k)}.$$

Note that the constant C above is chosen so that $d_{\mathcal{AC}}$ satisfies the triangle inequality. We equip the space $\mathbf{AC}((0,\infty))$ with the distance

$$d_{\mathbf{AC}}((\gamma_1,t_1),(\gamma_2,t_2))=d_{\mathcal{AC}}(\gamma_1,\gamma_2)+|t_2-t_1|$$

for $(\gamma_i,t_i)\in\mathcal{AC}(a_i)\times[0,a_i]$, $i=1,2$. It is not difficult to check that the spaces $\mathcal{AC}((0,\infty))$ and $\mathbf{AC}((0,\infty))$ with these distances are complete separable metric spaces and that the subsets $\mathcal{AC}(J)\subset\mathcal{AC}((0,\infty))$ and $\mathbf{AC}(J)\subset\mathbf{AC}((0,\infty))$ for a subset $J\subset(0,\infty)$ is compact if and only if J is compact.

From Corollary 3.3, the iterate F_*^n of the mapping $F_*: \mathcal{A}((0,\infty))\to\mathcal{A}((0,\infty))$ (resp. $F_*: \mathbf{A}((0,\infty))\to\mathbf{A}((0,\infty))$) for any $n\geqslant n_g$ carries the subset $\mathcal{AC}((0,\infty))$ (resp. $\mathbf{AC}((0,\infty))$) into itself. Further we have, for any $n\geqslant n_g$ and $a>0$,

$$F_*^n(\mathcal{AC}([a,\infty)))\subset\mathcal{AC}([a\exp(\lambda_g n-c_g),\infty)) \tag{12}$$

and

$$F_*^n(\mathbf{AC}([a,\infty)))\subset\mathbf{AC}([a\exp(\lambda_g n-c_g),\infty)). \tag{13}$$

We define the mapping $\Pi: \mathbf{A}((0,\infty))\to M$ and $\pi: \mathbf{A}((0,\infty))\to\mathcal{AC}((0,\infty))$ by $\Pi(\gamma,t)=\gamma(t)$ and $\pi(\gamma,t)=\gamma$. Obviously we have the commutative relations

$$\begin{array}{ccc} \mathbf{A}((0,\infty)) & \xrightarrow{F_*} & \mathbf{A}((0,\infty)) \\ \Pi\downarrow & & \Pi\downarrow \\ M & \xrightarrow{F} & M \end{array} \quad\text{and}\quad \begin{array}{ccc} \mathbf{A}((0,\infty)) & \xrightarrow{F_*} & \mathbf{A}((0,\infty)) \\ \pi\downarrow & & \pi\downarrow \\ \mathcal{A}((0,\infty)) & \xrightarrow{F_*} & \mathcal{A}((0,\infty)). \end{array} \tag{14}$$

3.5. Admissible measures

In this subsection, we are going to introduce the notion of admissible measure. First we introduce this notion in a simple case. Let $\gamma: [0,a]\to M$ be an admissible curve and, for $n\geqslant 1$, let $p_n: [0,a]\to[0,a_n]$ be the unique diffeomorphism that satisfies $p_n(0)=0$ and $dp_n(t)/dt=D_*F^n(\gamma(t),\gamma'(t))$ for $t\in[0,a]$. Since mappings $F\in\mathcal{U}$ act on the admissible curves as uniformly expanding mappings with uniformly bounded distortion, a standard argument on the iterations of uniformly expanding mappings gives the following result:

LEMMA 3.4. *The mapping p_n satisfies $dp_n(t)/dt\geqslant\exp(\lambda_g n-c_g)$ and*

$$\left|\log\frac{dp_n}{dt}(s)-\log\frac{dp_n}{dt}(s')\right|\leqslant C_g|p_n(s)-p_n(s')|\quad \textit{for } s,s'\in[0,a],$$

where C_g is the kind of constant that we mentioned in §3.3 and, in particular, does not depend on the mapping $F\in\mathcal{U}$, the admissible curve γ nor $n\geqslant n_g$.

We say that a measure μ on an interval $I\subset\mathbf{R}$ has *Lipschitz logarithmic density with constant L* if μ is written in the form $\mu=\varphi\,\mathbf{m_R}|_I$, where $\varphi: I\to\mathbf{R}$ is a positive-valued function satisfying

$$|\log\varphi(s)-\log\varphi(s')|\leqslant L|s-s'|\quad\text{for any } t,s\in I,$$

and $\mathbf{m_R}|_I$ is the restriction of the Lebesgue measure on $\mathbf{R}$ to I. Note that the sum (or integration) of measures on an interval I having Lipschitz logarithmic density with constant L again has the same property. From the lemma above, we can obtain a corollary:

COROLLARY 3.5. *There is a positive constant L_g such that, if a measure μ on $[0,a]$ has Lipschitz logarithmic density with constant L_g, then so does the measure $\mu\circ p_n^{-1}$ on $[0,a_n]$ for any $n\geqslant n_g$, any $F\in\mathcal{U}$ and any admissible curve $\gamma:[0,a]\to M$.*

We henceforth fix the constant L_g for which the claim of Corollary 3.5 holds. And we say that a measure ν on M is an admissible measure on an admissible curve $\gamma:[0,a]\to M$ if $\nu=\mu\circ\gamma^{-1}$ for a measure μ on $[0,a]$ that has Lipschitz logarithmic density with constant L_g. The following corollary is a consequence of Corollary 3.5:

COROLLARY 3.6. *If a measure ν is an admissible measure on an admissible curve $\gamma:[0,a]\to M$, then, for $n\geqslant n_g$ and $F\in\mathcal{U}$, the measure $\nu\circ F^{-n}$ is an admissible measure on the admissible curve $F_*^n\gamma$.*

We have introduced the notion of admissible measure on a single curve and seen that the iterates of mappings $F\in\mathcal{U}$ preserve such a class of measures. Now we are going to introduce more general definitions. Let $\Xi_{\mathbf{AC}}$ be the measurable partition of the space $\mathbf{AC}((0,\infty))$ into the subsets $\{\gamma\}\times[0,a]$ for $a>0$ and $\gamma\in\mathcal{AC}(a)$. In other words, we put $\Xi_{\mathbf{AC}}=\pi^{-1}\varepsilon$, where ε is the measurable partition of $\mathcal{AC}((0,\infty))$ into individual points and π is the mapping defined at the end of the last subsection. On each element $\xi=\{\gamma\}\times[0,a]$ of the partition $\Xi_{\mathbf{AC}}$, we consider the measure $\mathbf{m}_\xi$ that corresponds to the Lebesgue measure on $[0,a]$ through the bijection $(\gamma,t)\mapsto t$. For a Borel finite measure $\tilde{\mu}$ on $\mathbf{AC}((0,\infty))$, let $\{\tilde{\mu}_\xi\}_{\xi\in\Xi_{\mathbf{AC}}}$ be the conditional measures with respect to the partition $\Xi_{\mathbf{AC}}$. We put the following two definitions:

Definition. A Borel finite measure $\tilde{\mu}$ on $\mathbf{AC}((0,\infty))$ is said to be an *admissible measure* if the conditional measures $\{\tilde{\mu}_\xi\}_{\xi\in\Xi_{\mathbf{AC}}}$ have Lipschitz logarithmic density with constant L_g, $\tilde{\mu}$-almost everywhere.

Definition. A Borel finite measure μ on M is said to *have an admissible lift* if there exists an admissible measure $\tilde{\mu}$ on $\mathbf{AC}((0,\infty))$ such that $\tilde{\mu}\circ\Pi^{-1}=\mu$. The measure $\tilde{\mu}$ is said to be an *admissible lift* of the measure μ.

For a subset $J\subset(0,\infty)$, let $\mathbf{AM}(J)$ be the set of admissible measures that is supported on $\mathbf{AC}(J)$, and $\mathcal{AM}(J)$ the set of measures on M that have admissible lifts contained in $\mathbf{AM}(J)$. Then we have the following results:

LEMMA 3.7. *If a measure $\tilde{\mu}$ belongs to $\mathbf{AM}([a,\infty))$ for some $a\geqslant 0$ and if $F\in\mathcal{U}$, then $\tilde{\mu}\circ F_*^{-n}$ belongs to $\mathbf{AM}([a',\infty))$ for $n\geqslant n_g$, where $a'=a\exp(\lambda_g n-c_g)>a$.*

Proof. The conditional measures of the measure $\tilde{\mu}\circ F_*^{-n}$ with respect to the partition $\Xi_{\mathbf{AC}}$ are given as integrations of the images of the conditional measures $\{\tilde{\mu}_\xi\}_{\xi\in\Xi_{\mathbf{AC}}}$ under the mapping F_*^n. From Corollary 3.5 and the fact noted just above it, they have Lipschitz logarithmic density with constant L_g. Hence $\tilde{\mu}\circ F_*^{-n}$ is an admissible measure. The claim of the lemma follows from this and (13). □

COROLLARY 3.8. *If $\mu\in\mathcal{AM}([a,\infty))$ for some $a>0$ and if $F\in\mathcal{U}$, then the measure $\mu\circ F^{-n}$ belongs to $\mathcal{AM}([a',\infty))$ for $n\geqslant n_g$, where $a'=a\exp(\lambda_g n-c_g)>a$. In particular, if an invariant measure for $F\in\mathcal{U}$ has an admissible lift, it belongs to $\mathcal{AM}([a,\infty))$ for any $a>0$.*

LEMMA 3.9. *The subset $\mathbf{AM}(J)$ for a closed subset $J\subset(0,\infty)$ is closed in the space of Borel finite measures on $\mathbf{AC}((0,\infty))$.*

Proof. For a real number ε, we define the mapping T_ε from $\mathcal{AC}((0,\infty))\times\mathbf{R}$ to itself by $T_\varepsilon(\gamma,t)=(\gamma,t+\varepsilon)$. Then a measure $\tilde{\mu}$ on $\mathbf{AC}((0,\infty))\subset\mathcal{AC}((0,\infty))\times\mathbf{R}$ is admissible if and only if it satisfies

$$\int_{\mathbf{AC}((0,\infty))\cap T_\varepsilon^{-1}(\mathbf{AC}((0,\infty)))}\varphi\circ T_\varepsilon^{-1}\,d\tilde{\mu}\leqslant\exp(L_g|\varepsilon|)\int_{\mathbf{AC}((0,\infty))}\varphi\,d\tilde{\mu}$$

for any non-negative-valued continuous function φ on $\mathcal{AC}((0,\infty))\times\mathbf{R}$ and for any $\varepsilon>0$. For each non-negative-valued continuous function φ on $\mathcal{AC}((0,\infty))\times\mathbf{R}$ and $\varepsilon\in\mathbf{R}$, the set of Borel measures $\tilde{\mu}$ that satisfy the inequality above and that are supported on $\mathbf{AC}(J)$ is a closed subset in the space of Borel finite measures on $\mathbf{AC}((0,\infty))$. Hence so is their intersection, $\mathbf{AM}(J)$. □

LEMMA 3.10. *$\mathcal{AM}([a,\infty))=\mathcal{AM}([a,2a])$ for $a>0$.*

Proof. For $a>0$, let $\Delta_a\colon\mathbf{AC}([a,\infty))\to\mathbf{AC}([a,2a])$ be the mapping that brings an element $(\gamma,t)\in\mathcal{AC}(b)\times[0,b]$ to

$$\Delta_a((\gamma,t))=(\gamma|_{[m(t),m(t)+b/n]},t-m(t))\in\mathcal{AC}(b/n)\times[0,b/n],\tag{15}$$

where $n=[b/a]$ and $m(t)=[tn/b]\,b/n$. Then we have $\Pi\circ\Delta_a=\Pi$, and for any $\tilde{\mu}\in\mathbf{AM}([a,\infty))$, the image $\tilde{\mu}\circ\Delta_a^{-1}$ belongs to $\mathbf{AM}([a,2a])$. Thus we obtain the lemma. □

From the lemma above and Lemma 3.9, a corollary follows:

COROLLARY 3.11. *The set $\mathcal{AM}([a,\infty))$ for $a>0$ is a compact subset in the space of Borel finite measures on M. In particular, for a mapping $F\in\mathcal{U}$, the subset of F-invariant Borel probability measures that have admissible lifts is compact.*

Suppose that P is a small parallelogram on the torus $\mathbf{T}$ whose center z belongs to M and two of whose sides are parallel to the unstable subspace $\mathbf{E}^u(z)$. Then the restriction of the Lebesgue measure $\mathbf{m}$ to P has an admissible lift, provided that P is sufficiently small. Moreover any linear combination of such measures has admissible lifts. Thus we obtain the following result:

LEMMA 3.12. *For any Borel finite measure ν on M that is absolutely continuous with respect to the Lebesgue measure* $\mathbf{m}$, *there exist a sequence $b_n \to +0$ and measures $\nu_n \in \mathcal{AM}([b_n, \infty))$ such that $|\nu - \nu_n| \to 0$ as $n \to \infty$. Further we can take the measures ν_n so that the densities $d\nu_n/d\mathbf{m}$ are square integrable.*

The next lemma is a consequence of the last two lemmas and Corollary 3.8.

LEMMA 3.13. *Let F be a mapping in $\mathcal{U}$ and ν a probability measure on M that is absolutely continuous with respect to the Lebesgue measure* $\mathbf{m}$. *Then any limit point of the sequence $n^{-1} \sum_{i=0}^{n-1} \nu \circ F^{-i}$ is contained in $\mathcal{AM}([a, \infty))$ for any $a>0$. In particular, physical measures for F are contained in $\mathcal{AM}([a, \infty))$ for any $a>0$.*

Finally we prove another lemma.

LEMMA 3.14. *Let F be a mapping in $\mathcal{U}$. If an F-invariant Borel probability measure has an admissible lift, so do its ergodic components.*

Proof. From Corollary 3.11, it is enough to show the following claim: If an F-invariant measure μ that has an admissible lift splits into two non-trivial F-invariant measures μ_1 and μ_2 that are totally singular with respect to each other, then the measures μ_1 and μ_2 have admissible lifts. We are going to show this claim. From Corollary 3.8, we can take an admissible lift $\tilde{\mu}$ of μ that is supported on $\mathbf{AC}([1, \infty))$. Consider the mapping $G = \Delta_1 \circ F_*^{n_g} : \mathbf{AC}([1, \infty)) \to \mathbf{AC}([1, 2])$, where Δ_1 is the mapping defined by (15). Then the measure $\tilde{\mu} \circ G^{-1}$ is an admissible lift of μ. Replacing $\tilde{\mu}$ by $\tilde{\mu} \circ G^{-1}$, we can assume that $\tilde{\mu}$ is supported on $\mathbf{AC}([1, 2])$. From the assumption of the claim, we can take an F-invariant Borel subset $X \subset M$ such that $\mu_1(M \setminus X) = \mu_2(X) = 0$. Then, by the relation $F^{n_g} \circ \Pi = \Pi \circ G$, the set $\widetilde{X} := \Pi^{-1}(X)$ is G-invariant. Below we prove that $\widetilde{X}$ is a $\Xi_{\mathbf{AC}}$-set, that is, a union of elements of the partition $\Xi_{\mathbf{AC}}$, modulo null subsets with respect to $\tilde{\mu}$. This implies the claim above because the restriction of the measure $\tilde{\mu}$ to $\widetilde{X}$ is an admissible lift of μ_1.

Put $\Xi_1 = \Xi_{\mathbf{AC}}$ and define the sequence Ξ_n, $n = 1, 2, \dots$, inductively by the relation $\Xi_{n+1} = G^{-1}(\Xi_n) \vee \Xi_1$. Then Ξ_n is increasing with respect to n and the limit $\bigvee_{n=1}^{\infty} \Xi_n$ is the measurable partition into individual points. Thus the conditional expectation $E(\widetilde{X} | \Xi_n)$ with respect to $\tilde{\mu}$ converges to the indicator function of $\widetilde{X}$ as $n \to \infty$, $\tilde{\mu}$-almost everywhere. Note that the restriction of G^n to each element of the partition Ξ_n is a bijection onto an element of Ξ_1, and its distortion is uniformly bounded. Hence, using the assumption that $\tilde{\mu}$ is an admissible measure and the invariance of $\widetilde{X}$, we can see that the conditional expectation $E(\widetilde{X} | \Xi_1)$ equals the indicator function of $\widetilde{X}$, or $\widetilde{X}$ is a $\Xi_{\mathbf{AC}}$-set modulo null subsets with respect to $\tilde{\mu}$. □

3.6. The no flat contact condition

In this subsection, we consider the influence of the critical points on ergodic behavior of partially hyperbolic endomorphisms. We first explain a problem that the critical points may cause. And then we give a mild condition on the mappings in $\mathcal{U}$, *the no flat contact condition*, which allows us to avoid that problem. In the last part of this paper, we will prove that this condition holds for almost all partially hyperbolic endomorphisms in $\mathcal{U}$.

Let us consider a mapping $F\in\mathcal{U}$. Let $\chi_c(z;F)<\chi_u(z;F)$ be the Lyapunov exponents at $z\in M$. For a Borel finite measure μ on M, we define

$$\chi_c(\mu;F)=\frac{1}{|\mu|}\int \log\|DF|_{E^c(z)}\|\,d\mu(z)\quad\text{and}\quad \chi_u(\mu;F)=\frac{1}{|\mu|}\int\log\frac{|\det DF_z|}{\|DF|_{E^c(z)}\|}\,d\mu(z).$$

These are called the central and unstable Lyapunov exponent of μ, respectively. For an invariant probability measure μ for F, we have

$$\chi_c(\mu;F)=\int\chi_c(z)\,d\mu(z)\quad\text{and}\quad\chi_u(\mu;F)=\int\chi_u(z)\,d\mu(z).$$

Further, if μ is an ergodic invariant measure for $F\in\mathcal{U}$, the Lyapunov exponents $\chi_c(z;F)$ and $\chi_u(z;F)$ take constant values $\chi_c(\mu;F)$ and $\chi_u(\mu;F)$ at μ-almost every point z, respectively.

Let μ_n, $n=1,2,\dots$, be a sequence of ergodic invariant probability measures for F that have admissible lifts, and suppose that μ_n converges weakly to some measure μ_∞ as $n\to\infty$. Then μ_∞ has an admissible lift from Corollary 3.11. It is not difficult to see that the Lyapunov exponent $\chi_u(\mu_n;F)$ always converges to $\chi_u(\mu_\infty;F)$. However, for the central Lyapunov exponent, we only have the inequality

$$\limsup_{i\to\infty}\chi_c(\mu_n;F)\leqslant\chi_c(\mu_\infty;F)\tag{16}$$

when F has critical points, because the function $\log\|DF|_{E^c(z)}\|$ is not continuous at the critical points. Though the strict inequality in (16) is not likely to hold often, we cannot avoid such cases in general. And, once the strict inequality holds, the ergodic behavior of F can be wild by the influence of the critical point.

Remark. It is not easy to construct examples in which the strict inequality (16) holds. For example, consider the direct product of the quadratic mappings given in the paper [8] and an angle-multiplying mapping $\theta\mapsto d\theta$ on the circle.

Remark. We could consider a similar but more general problem: Suppose that a point $z\in M$ is generic for an invariant probability measure μ, that is, the sequence $n^{-1}\sum_{i=0}^{n-1}\delta_{F^i(z)}$ converges to μ as $n\to\infty$. The problem is that the strict inequality

$\chi_c(z;F)<\chi_c(\mu;F)$ can hold, though the equality $\chi_u(z;F)=\chi_u(\mu;F)$ always holds. (If we did not assume the mapping F to be partially hyperbolic, these relations would be looser.) We may call this kind of problems *Lyapunov irregularity*, as this is the case where the so-called Lyapunov regularity condition [13] does not hold.

In order to avoid the irregularity described above, we introduce a mild condition:

Definition. We say that a mapping $F\in\mathcal{U}$ satisfies *the no flat contact condition* if there exist positive constants $C=C(F)$, $n_0=n_0(F)\geqslant n_g$ and $\beta=\beta(F)$ such that, for any admissible curve $\gamma\in\mathcal{AC}(a)$ with $a>0$, $n\geqslant n_0$ and $\varepsilon>0$,

$$\mathbf{m}_{\mathbf{R}}(\{t\in[0,a]\mid d(F^n(\gamma(t)),\mathcal{C}(F))<\varepsilon\})<C\varepsilon^{\beta}\max\{a,1\},$$

where $\mathbf{m}_{\mathbf{R}}$ is the Lebesgue measure on $\mathbf{R}$. If F has no critical points, we say that $d(z,\mathcal{C}(F))=1$ for $z\in M$ and that F satisfies the no flat contact condition.

Remark. The definition above is motivated by the argument in a paper of Viana [25], in which the condition as above for $\beta=\frac{1}{2}$ is considered.

Below we give simple consequences of the no flat contact condition. For $F\in\mathcal{U}$ and $z\in M$, we define

$$L(z;F):=\log\min_{v\in\mathbf{S}^u(z)}|D^*F(z,v)|\in\mathbf{R}\cup\{-\infty\}. \tag{17}$$

This function is continuous outside the critical set $\mathcal{C}(F)$ and satisfies

$$L(z;F)\geqslant\log d(z,\mathcal{C}(F))-C_g$$

from (5), provided that $\mathcal{C}(F)\neq\varnothing$. Thus we get the following lemma:

LEMMA 3.15. *Suppose that $F\in\mathcal{U}$ satisfies the no flat contact condition and let $n_0=n_0(F)$ be as in the condition. For any $\delta>0$ and $a>0$, we can choose a positive number $h=h(\delta,a;F)$ such that*

$$\int\min\{0,L(z;F)+h\}\,d(\mu\circ F^{-n})(z)\geqslant-\delta|\mu|$$

for any $\mu\in\mathcal{AM}([a,\infty))$ and $n\geqslant n_0$.

Using the inequality $\log\|DF|_{E^c(z)}\|\geqslant L(z;F)-C_g$, which follows from (5), together with Lemma 3.15, Corollary 3.8 and Corollary 3.11, we can obtain the following corollary:

COROLLARY 3.16. *Suppose that a mapping $F\in\mathcal{U}$ satisfies the no flat contact condition. Then the central Lyapunov exponent $\chi_c(\mu;F)$, considered as a function on the space of F-invariant probability measures having admissible lifts, is continuous and, in particular, uniformly bounded away from $-\infty$.*

This corollary implies that the irregularity of the central Lyapunov exponent we mentioned does not take place under the no flat contact condition.

3.7. Multiplicity of tangencies between the images of the unstable cones

By an iterate of a mapping $F\in\mathcal{U}$, the unstable cones $\mathbf{S}^u(z)$ at many points z will be brought to one point, and some pairs of their images may be tangent, that is, have non-empty intersection. (Recall that $\mathbf{S}^u(z)$ does not contain the origin $\mathbf{0}$.) In this subsection, we introduce quantities that measure the multiplicity of such tangencies and then formulate a condition, *the transversality condition on unstable cones*, for mappings in $\mathcal{U}$.

We introduce analogues of the so-called Pesin subsets. Let $\chi=\{\chi_c^-,\chi_c^+,\chi_u^-,\chi_u^+\}$ be a quadruple of real numbers that satisfy

$$\chi_c^-<\chi_c^+<\chi_u^-<\chi_u^+. \tag{18}$$

Let ε be a small positive number. For a mapping $F\in\mathcal{U}$, an integer $n>0$ and a real number $k>0$, we define a closed subset $\Lambda(\chi,\varepsilon,k,n;F)$ of M as the set of all points $z\in M$ that satisfy

$$\chi_c^-(j-i)-\varepsilon(n-j)-k\leqslant\log|D^*F^{j-i}(v)|\leqslant\chi_c^+(j-i)+\varepsilon(n-j)+k$$

and

$$\chi_u^-(j-i)-\varepsilon(n-j)-k\leqslant\log D_*F^{j-i}(v)\leqslant\chi_u^+(j-i)+\varepsilon(n-j)+k$$

for any $0\leqslant i<j\leqslant n$ and $v\in\mathbf{S}^u(F^i(z))$. Applying the standard argument in the Pesin theory [16], [18] to the inverse limit system, we can show the following result:

LEMMA 3.17. *If μ is an invariant probability measure for $F\in\mathcal{U}$ and if*

$$\chi_c^-<\chi_c(z;F)<\chi_c^+\quad\text{and}\quad\chi_u^-<\chi_u(z;F)<\chi_u^+\quad\text{for }\mu\text{-almost every }z,$$

then $\lim_{k\to\infty}\liminf_{n\to\infty}\mu(\Lambda(\chi,\varepsilon,k,n;F))=1$.

Note that we have

$$\Lambda(\chi,\varepsilon,k,n;F)\subset\Lambda(\chi,\varepsilon',k',n;F)\qquad\text{if } k\leqslant k' \text{ and } \varepsilon\leqslant\varepsilon', \tag{19}$$

$$F^i(\Lambda(\chi,\varepsilon,k,n;F))\subset\Lambda(\chi,\varepsilon,k,n-i;F)\qquad\text{for } 0\leqslant i<n, \tag{20}$$

$$\Lambda(\chi,\varepsilon,k,n;F)\subset\Lambda(\chi,\varepsilon,k+\varepsilon i,n-i;F)\quad\text{for } 0\leqslant i<n. \tag{21}$$

By (4), we can take a constant H_g such that

$$\angle(DF^n(u),DF^n(v))<H_g\frac{|D^*F^n(z,w)|}{D_*F^n(z,w)}\leqslant H_g\exp((\chi_c^+-\chi_u^-)n+2k) \tag{22}$$

for any $z\in\Lambda(\chi,\varepsilon,k,n;F)$ and $u,v,w\in\mathbf{S}^u(z)$. For $z\in M$, let $\mathcal{E}(z;\chi,\varepsilon,k,n;F)$ be the set of all pairs (w,w') of points in $F^{-n}(z)\cap\Lambda(\chi,\varepsilon,k,n;F)$ such that

$$\angle(DF^n(\mathbf{E}^u(w')),DF^n(\mathbf{E}^u(w)))\leqslant 5H_g\exp((\chi_c^+-\chi_u^-)n+2k). \tag{23}$$

Note that, if a pair (w,w') of points in $F^{-n}(z)\cap\Lambda(\chi,\varepsilon,k,n;F)$ does *not* belong to $\mathcal{E}(z;\chi,\varepsilon,k,n;F)$, we have

$$\angle(DF^n(u),DF^n(u'))>3H_g\exp((\chi_c^+-\chi_u^-)n+2k) \tag{24}$$

for any $u\in\mathbf{S}(w)$ and $u'\in\mathbf{S}(w')$, from (22).

As a measure for the multiplicity of tangencies, we consider the number

$$\mathbf{N}(\chi,\varepsilon,k,n;F)=\max_{z\in M}\max_{w\in F^{-n}(z)\cap\Lambda(\chi,\varepsilon,k,n;F)}\#\{w'\mid(w,w')\in\mathcal{E}(z;\chi,\varepsilon,k,n;F)\}.$$

This is increasing with respect to k and ε.

Definition. Let $\mathbf{X}=\{\chi(l)\}_{l=1}^{l_0}$ be a finite collection of quadruples of numbers $\chi(l)=\{\chi_c^-(l),\chi_c^+(l),\chi_u^-(l),\chi_u^+(l)\}$ that satisfy (18). We say that a mapping $F\in\mathcal{U}$ satisfies *the transversality condition on unstable cones for* $\mathbf{X}$ if

$$\lim_{\varepsilon\to+0}\lim_{k\to\infty}\liminf_{n\to\infty}\max\left\{\frac{\log\mathbf{N}(\chi(l),\varepsilon,k,n;F)}{n(\chi_c^-(l)+\chi_u^-(l)-\chi_c^\Delta(l)-\chi_u^\Delta(l))}\;\middle|\;1\leqslant l\leqslant l_0\right\}<1,$$

where $\chi_c^\Delta(l)=\chi_c^+(l)-\chi_c^-(l)$ and $\chi_u^\Delta(l)=\chi_u^+(l)-\chi_u^-(l)$.

Remark. We will consider only the case where $\chi_c^-(l)+\chi_u^-(l)-\chi_c^\Delta(l)-\chi_u^\Delta(l)>0$.

3.8. Measures on the space of mappings

In this subsection, we give some additional arguments concerning measures on the space of mappings. Recall that $\tau_\psi\colon C^r(M,\mathbf{R}^2)\to C^r(M,\mathbf{R}^2)$ is the translation by $\psi\in C^r(M,\mathbf{R}^2)$, that is, $\tau_\psi(\varphi)=\varphi+\psi$. For an integer $s\geqslant 0$ and a positive number $d>0$, we put

$$\mathbf{D}^s(d)=\{G\in C^s(M,\mathbf{R}^2)\mid\|G\|_{C^s}\leqslant d\}. \tag{25}$$

The following lemma gives measures on $C^r(M,\mathbf{R}^2)$ with nice properties:

LEMMA 3.18. *For an integer $s\geqslant 3$, there exists a Borel probability measure $\mathcal{M}_s$ on $C^{s-3}(M,\mathbf{R}^2)$ such that*

(1) *$\mathcal{M}_s$ is quasi-invariant along $C^{s-1}(M,\mathbf{R}^2)$;*

(2) *there exists a positive constant $\varrho=\varrho_s(d)$ for any $d>0$ such that*

$$\frac{1}{2}\leqslant\frac{d(\mathcal{M}_s\circ\tau_\psi^{-1})}{d\mathcal{M}_s}\leqslant 2\quad \mathcal{M}_s\text{-almost everywhere on }\mathbf{D}^{s-3}(d)$$

for any $\psi\in C^s(M,\mathbf{R}^2)$ with $\|\psi\|_{C^s}<\varrho$.

We will give the proof of Lemma 3.18 in the appendix at the end of this paper. This is on the one hand because the lemma itself has nothing to do with dynamical systems, and on the other hand because the proof is merely a combination of some results in probability theory.

Henceforth, we fix the measures $\mathcal{M}_s$ for $s\geqslant 3$ in Lemma 3.18. Note that the measure $\mathcal{M}_s$ belongs to $\mathcal{Q}^r_{s-1}$ when $s\geqslant r+3$.

LEMMA 3.19. *Suppose that $s\geqslant r+3$. If a Borel subset X in $C^r(M,M)$ is shy with respect to the measure $\mathcal{M}_{s+2}$, then X is timid for the class $\mathcal{Q}^r_{s-1}$ of measures.*

Proof. Take an arbitrary measure $\mathcal{N}$ in $\mathcal{Q}^r_{s-1}$. The measure $\mathcal{M}_{s+2}$ is supported on the space $C^{s-1}(M,\mathbf{R}^2)$, along which $\mathcal{N}$ is quasi-invariant. Hence the convolution $\mathcal{N}*\mathcal{M}_{s+2}$ is equivalent to $\mathcal{N}$. From the assumption, we have

$$\mathcal{N}*\mathcal{M}_{s+2}(\Phi_G^{-1}(X))=\int \mathcal{M}_{s+2}\circ\tau_\psi^{-1}(\Phi_G^{-1}(X))\,d\mathcal{N}(\psi)=\int \mathcal{M}_{s+2}(\Phi_{G+\psi}^{-1}(X))\,d\mathcal{N}(\psi)=0$$

for any $G\in C^r(M,\mathbf{T})$. Thus X is shy with respect to $\mathcal{N}$. □

In order to evaluate subsets in $C^r(M,\mathbf{T})$ with respect to the measures $\mathcal{M}_s$, we will use the following lemma:

LEMMA 3.20. *Let $s\geqslant r+3$ and $d>0$. Suppose that mappings $\psi_i\in C^s(M,\mathbf{R}^2)$ and positive numbers T_i for $1\leqslant i\leqslant m$ satisfy*

$$\sup_{|t_i|\leqslant T_i}\left\|\sum_{i=1}^m t_i\psi_i\right\|_{C^s}\leqslant \varrho_s(d), \tag{26}$$

where $\varrho_s(d)$ is as in Lemma 3.18. If a Borel subset X in $C^r(M,\mathbf{T})$ satisfies, for some $\beta>0$, that

$$\mathbf{m}_{\mathbf{R}^m}\left(\left\{\{t_i\}_{i=1}^m\in\prod_{i=1}^m[-T_i,T_i]\;\middle|\;\varphi+\sum_{i=1}^m t_i\psi_i\in X\right\}\right)<\beta\prod_{i=1}^m 2T_i \tag{27}$$

for every $\varphi\in X$, then we have

$$\mathcal{M}_s(\Phi_G^{-1}(X)\cap\mathbf{D}^{s-3}(d))\leqslant 2^{m+1}\beta\mathcal{M}_s(\Phi_G^{-1}(Y))\leqslant 2^{m+1}\beta$$

for any $G\in C^r(M,\mathbf{T})$, where

$$Y=\left\{\psi+\sum_{i=1}^m t_i\psi_i\;\middle|\;\psi\in X \text{ and } |t_i|\leqslant\frac{T_i}{2}\right\}.$$

Proof. Put $Z=\Phi_G^{-1}(X)\cap\mathbf{D}^{s-3}(d)$ and let $\mathbf{1}_Z$ be the indicator (characteristic) function of Z. Using the Fubini theorem and the properties of $\mathcal{M}_s$, we get

$$\int \mathbf{m}_{\mathbf{R}^m}\left(\left\{\mathbf{t}\in\mathbf{R}^m \;\middle|\; |t_i|\leqslant\frac{T_i}{2} \text{ and } \tilde{\psi}+\sum_{i=1}^m t_i\psi_i\in Z\right\}\right)d\mathcal{M}_s(\tilde{\psi})$$
$$=\int_{\{\mathbf{t}\,|\,|t_i|<\frac{1}{2}T_i\}}\left(\int \mathbf{1}_Z\left(\tilde{\psi}+\sum_{i=1}^m t_i\psi_i\right)d\mathcal{M}_s(\tilde{\psi})\right)d\mathbf{m}_{\mathbf{R}^m}(\mathbf{t})$$
$$=\int_{\{\mathbf{t}\,|\,|t_i|<\frac{1}{2}T_i\}}\mathcal{M}_s\left(Z-\sum_{i=1}^m t_i\psi_i\right)d\mathbf{m}_{\mathbf{R}^m}(\mathbf{t})$$
$$\geqslant\frac{1}{2}\mathcal{M}_s(Z)\prod_{i=1}^m T_i.$$

The integrand of the integral on the first line is positive only if $\tilde{\psi}$ belongs to $\Phi_G^{-1}(Y)$ and bounded by $\beta\prod_{i=1}^m 2T_i$ from the assumption (27). Thus we obtain the lemma. □

3.9. The plan of the proof of the main theorems

Now we can describe the plan of the proof of the main results, Theorems 2.1 and 2.2, more concretely by using the terminology introduced in the preceding subsections. We split the proof into two parts. In the former part, which will be carried out in §§4–6, we study ergodic properties of partially hyperbolic endomorphisms in $\mathcal{U}$ that satisfy the no flat contact condition and the transversality condition on unstable cones for some finite collection of quadruples. The conclusion in this part is the following theorem. For a finite or countable collection $\mathbf{X}=\{\chi(l)\}_{l\in L}$ of quadruples $\chi(l)=\{\chi_c^-(l),\chi_c^+(l)),\chi_u^-(l),\chi_u^+(l)\}$ that satisfy the condition (18), let $|\mathbf{X}|$ be the union of the open rectangles $(\chi_c^-(l),\chi_c^+(l))\times(\chi_u^-(l),\chi_u^+(l))$ over $l\in L$.

THEOREM 3.21. *Let* $\mathbf{X}$ *be a finite collection of quadruples that satisfy* (18),

$$\chi_c^-<0, \tag{28}$$
$$\chi_c^-+\chi_u^->(\chi_c^+-\chi_c^-)+(\chi_u^+-\chi_u^-)>0 \tag{29}$$

and also

$$\{0\}\times[\lambda_g,\Lambda_g]\subset|\mathbf{X}|\subset(-2\Lambda_g,2\Lambda_g)\times(0,2\Lambda_g). \tag{30}$$

Suppose that a mapping F *in* $\mathcal{U}$ *satisfies the no flat contact condition and the transversality condition on unstable cones for* $\mathbf{X}$. *Then* F *admits a finite collection of ergodic physical measures whose union of basins of attraction has total Lebesgue measure on* M. *In addition, if an ergodic physical measure* μ *for* F *satisfies either* $(\chi_c(\mu;F),\chi_u(\mu;F))\in|\mathbf{X}|$ *or* $\chi_c(\mu;F)>0$, *then* μ *is absolutely continuous with respect to the Lebesgue measure* $\mathbf{m}$.

In the latter part of the proof, which will be carried out in §§7 and 8, we show that the two conditions assumed on the mapping F in the theorem above hold for almost all partially hyperbolic endomorphisms in $\mathcal{U}$, provided that we choose the finite collection $\mathbf{X}$ of quadruples appropriately. On the one hand, we will prove the following theorem in §7. For a finite collection $\mathbf{X}$ of quadruples that satisfy (18), let $\mathcal{S}_1(\mathbf{X})$ be the set of mappings $F\in\mathcal{U}$ that does *not* satisfy the transversality condition on unstable cones for $\mathbf{X}$.

THEOREM 3.22. *There exists a countable collection* $\mathbf{X}=\{\chi(l)\}_{l=1}^{\infty}$ *of quadruples satisfying* (18), (28) *and* (29) *such that*

(a) $|\mathbf{X}|$ *contains the subset* $\{(x_c,x_u)\in\mathbf{R}^2\,|\,x_c+x_u>0,\ \lambda_g\leqslant x_u\leqslant\Lambda_g \text{ and } x_c\leqslant 0\}$;

(b) $|\mathbf{X}|$ *is contained in* $(-2\Lambda_g,2\Lambda_g)\times(0,2\Lambda_g)$;

(c) *the subset* $\mathcal{S}_1(\mathbf{X}')$ *for any finite subcollection* $\mathbf{X}'\subset\mathbf{X}$ *is shy with respect to the measures* $\mathcal{M}_s$ *for* $s\geqslant r+3$ *and is a meager subset in* $\mathcal{U}$ *in the sense of Baire's category argument.*

On the other hand, we will show the following theorem in §8. Let $\mathcal{S}_2$ be the set of mappings $F\in\mathcal{U}$ that does *not* satisfy the no flat contact condition.

THEOREM 3.23. *If an integer* $s\geqslant r+3$ *satisfies the condition* (3) *for some integer* $3\leqslant\nu\leqslant r-2$, *then the subset* $\mathcal{S}_2$ *is shy with respect to the measure* $\mathcal{M}_s$. *Moreover,* $\mathcal{S}_2$ *is contained in a closed nowhere dense subset in* $\mathcal{U}$, *provided that* $r\geqslant 19$.

It is easy to check that the three theorems above imply the main theorems: Consider a countable set of quadruples $\mathbf{X}=\{\chi(l)\}_{l=1}^{\infty}$ in Theorem 3.22 and put $\mathbf{X}_m=\{\chi(l)\}_{l=1}^{m}$ for $m>0$. Theorem 3.21 implies that the complement of $(\bigcup_{m=1}^{\infty}\mathcal{S}_1(\mathbf{X}_m))\cup\mathcal{S}_2$ in $\mathcal{U}$ is contained in $\mathcal{R}^r$. Thus the main theorems, Theorems 2.1 and 2.2, restricted to $\mathcal{U}$ follow from Theorems 3.22, 3.23 and Lemma 3.19. As we noted in §3.2, this is enough for the proof of the main theorems.

4. Hyperbolic physical measures

In this section, we study hyperbolic physical measures for partially hyperbolic endomorphisms. Throughout this section, *we consider a mapping F in $\mathcal{U}$ that satisfies the no flat contact condition.*

4.1. Physical measures with negative central exponent

In this subsection, we study physical measures whose central Lyapunov exponent is negative.

LEMMA 4.1. *If an ergodic probability measure μ with negative central Lyapunov exponent has an admissible lift, then it is a physical measure.*

Proof. The central Lyapunov exponent of the measure μ is bounded away from $-\infty$ by Corollary 3.16. From Oseledets's theorem and the assumption that μ has an admissible lift, we can find an admissible curve γ such that almost all points with respect to the smooth measure on it are forward Lyapunov regular for μ. According to the Pesin theory, the so-called Pesin's local stable manifold exists for all such points on γ. These local stable manifolds are transversal to γ and contained in the basin $\mathcal{B}(\mu)$ of μ. Further, the union of them has positive Lebesgue measure from absolute continuity of Pesin's local stable manifolds [18, §4]. Therefore μ is a physical measure. □

From this lemma and Lemma 3.14, we get the following result:

COROLLARY 4.2. *If an F-invariant probability measure μ has an admissible lift, it has at most countably many ergodic components with negative central Lyapunov exponent, each of which is a physical measure and absolutely continuous with respect to μ.*

The basin of an ergodic physical measure with negative central Lyapunov exponent may have empty interior, even though we ignore null subsets with respect to the Lebesgue measure $\mathbf{m}$. Nevertheless, we have the following lemmas:

LEMMA 4.3. *For an ergodic physical measure μ with negative central Lyapunov exponent, there is an open subset U with $\mu(U)=1$ such that, for a Borel finite measure ν that has an admissible lift, we have $\nu(\mathcal{B}(\mu))>0$ if and only if $\limsup_{n\to\infty}\nu\circ F^{-n}(U)>0$. In particular, if we assume ν to be F-invariant, we have $\nu(\mathcal{B}(\mu))>0$ if and only if $\nu(U)>0$.*

Proof. Recall the proof of Lemma 4.1. From absolute continuity of Pesin's local stable manifolds, there exists an open neighborhood U_z for μ-almost every point z such that, if an admissible curve $\gamma\colon[0,a]\to M$ with length $a>2$ satisfies $\gamma([1,a-1])\cap U_z\neq\varnothing$, the inverse image $\gamma^{-1}(\mathcal{B}(\mu))$ has positive Lebesgue measure. Let U be the union of such neighborhoods U_z. Then we obviously have $\mu(U)=1$. If $\limsup_{n\to\infty}\nu\circ F^{-n}(U)>0$ for a Borel finite measure ν that has an admissible lift, we have $\nu(\mathcal{B}(\mu))>0$ from the choice of U_z and Corollary 3.8. Conversely, if we have $\nu(\mathcal{B}(\mu))>0$, then

$$\limsup_{n\to\infty}\nu\circ F^{-n}(U)\geqslant\nu(\mathcal{B}(\mu))\,\mu(U)>0.$$ □

LEMMA 4.4. *Let μ_i, $i=1,2,\ldots$, be a sequence of mutually distinct F-invariant Borel probability measures each of which is ergodic and has an admissible lift. If μ_i converges to some measure μ_∞ as $i\to\infty$, we have $\chi_c(z;F)\geqslant0$ for μ_∞-almost every $z\in M$.*

Proof. From Corollary 3.11, μ_∞ has an admissible lift. If the conclusion of the lemma were not true, there should be an ergodic physical measure $\mu'_\infty \ll \mu_\infty$ with negative central Lyapunov exponent, from Corollary 4.2. Take the open set U in Lemma 4.3 for μ'_∞. On the one hand, $\mu'_\infty(U)=1$ and hence $\mu_\infty(U)>0$. On the other hand, since $\mu_i \neq \mu'_\infty$ except for one i at most, we should have $\mu_i(\mathcal{B}(\mu'_\infty))=0$ and hence $\mu_i(U)=0$. This contradicts the fact that μ_i converges to μ_∞. □

From this lemma and Corollary 3.16, the next corollary follows:

COROLLARY 4.5. *For any negative number $\chi<0$, there exist at most finitely many ergodic physical measures for F that satisfies $\chi_c(\mu;F)\leqslant\chi$.*

Finally we show a lemma:

LEMMA 4.6. *Let ν be a Borel finite measure that is absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$, and let μ be a limit point of the sequence of measures $n^{-1}\sum_{i=0}^{n-1}\nu\circ F^{-i}$, $n=1,2,\dots$. Then we have either*

(a) *$\chi_c(z;F)\geqslant 0$ for μ-almost every point $z\in M$, or*

(b) *there is an ergodic physical measure $\mu'\ll\mu$ with negative central Lyapunov exponent and $\nu(\mathcal{B}(\mu'))>0$.*

In particular, for a physical measure μ for F, we have either (a) *or*

(b′) *μ is ergodic and has negative central Lyapunov exponent.*

Proof. Suppose that (a) does not hold. Then, from Corollary 4.2, there exists an ergodic physical measure $\mu'\ll\mu$ with negative central Lyapunov exponent. Take the open set U in Lemma 4.3 for μ'. We have $\mu'(U)=1$ and hence $\mu(U)>0$. Thus

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\nu\circ F^{-i}(U)\geqslant\mu(U)>0.$$

Although the measure ν may not have an admissible lift, we can use the approximation in Lemma 3.12 to conclude that $\nu(\mathcal{B}(\mu'))>0$ from the property of U. □

4.2. Physical measures with positive central exponent

In this subsection, we investigate physical measures with positive central Lyapunov exponent. We shall prove the following three propositions:

PROPOSITION 4.7. *Any physical measure μ with positive central Lyapunov exponent is ergodic and absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$. Moreover, the basin $\mathcal{B}(\mu)$ is an open set modulo Lebesgue null subsets.*

PROPOSITION 4.8. *For any positive number $\chi>0$, there exist at most finitely many ergodic physical measures for F that satisfies $\chi_c(\mu;F)\geqslant\chi$.*

Let $\mathcal{B}^+(F)$ (resp. $\mathcal{B}^-(F)$) be the union of the basins of ergodic physical measures with positive (resp. negative) central Lyapunov exponent.

PROPOSITION 4.9. *Suppose that a Borel probability measure ν on M is absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$ and supported on the complement of $\mathcal{B}^-(F)\cup\mathcal{B}^+(F)$. If ν_∞ is a weak limit point of the sequence of measures $n^{-1}\sum_{j=0}^{n-1}\nu\circ F^{-j}$, $n=1,2,\ldots$, then we have $\chi_c(z;F)=0$ for ν_∞-almost every point z.*

We derive the propositions above from the following single proposition: Let $X(i)$, $i=1,2,\ldots$, be Borel subsets in M with positive Lebesgue measure. Let $\mathbf{m}_{X(i)}$ be the normalization of the restriction of the Lebesgue measure $\mathbf{m}$ to $X(i)$. For each $i\geqslant 1$, let $\mu_{i,\infty}$ be a weak limit point of the sequence $n^{-1}\sum_{j=0}^{n-1}\mathbf{m}_{X(i)}\circ F^{-j}$, $n=1,2,\ldots$. Assume that the sequence $\mu_{i,\infty}$ converges weakly to some measure μ_∞ as $i\to\infty$. Also assume that $\chi_c(\mu_\infty;F)>0$ and that $\chi_c(z;F)\geqslant 0$, μ_∞-almost everywhere.

PROPOSITION 4.10. *In the situation as above, there exist an ergodic physical measure $\nu_{i,\infty}$ and an open disk D_i in M for sufficiently large i such that*

(a) *$\nu_{i,\infty}\ll\mu_{i,\infty}$ and $\nu_{i,\infty}\ll\mathbf{m}$;*
(b) *$\chi_c(\nu_{i,\infty};F)>0$;*
(c) *the radius of D_i is positive and independent of i;*
(d) *$\nu_{i,\infty}(D_i)>0$ and $D_i\subset\mathcal{B}(\nu_{i,\infty})$ modulo Lebesgue null subsets.*

Below we prove Propositions 4.7, 4.8 and 4.9 using Proposition 4.10.

Proof of Proposition 4.7. Let μ be a physical measure such that $\chi_c(\mu;F)>0$. From Lemma 4.6, we have $\chi_c(z;F)\geqslant 0$ for μ-almost every point z. Apply Proposition 4.10 to the situation where $X(i):=\mathcal{B}(\mu)$ and $\mu_{i,\infty}=\mu_\infty=\mu$ for all $i\geqslant 1$. And let $\nu_{i,\infty}$ and D_i be those in the corresponding conclusion, which we can assume to be independent of i. Consider the open set $V=\bigcup_{n=0}^{\infty}F^{-n}(D_i)$. Then $\mathcal{B}(\nu_{i,\infty})=V$ modulo Lebesgue null subsets. Since $\nu_{i,\infty}(V)\geqslant\nu_{i,\infty}(D_i)>0$ and since $\nu_{i,\infty}\ll\mu$, we have $\mu(V)>0$. Hence

$$\mathbf{m}_{\mathcal{B}(\mu)}(\mathcal{B}(\nu_{i,\infty}))=\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mathbf{m}_{\mathcal{B}(\mu)}\circ F^{-i}(\mathcal{B}(\nu_{i,\infty}))\geqslant\mu(V)>0.$$

This implies $\mu=\nu_{i,\infty}$. We have proved Proposition 4.7. □

Proof of Proposition 4.8. Suppose that there exist infinitely many ergodic physical measures μ_i, $i=1,2,\ldots$, that satisfy $\chi_c(\mu_i;F)\geqslant\chi>0$. By taking a subsequence, we assume that μ_i converges to an invariant probability measure μ_∞ as $i\to\infty$. Then we have

$\chi_c(\mu_\infty; F) \geqslant \chi > 0$ from Corollaries 3.11 and 3.16. From Lemma 4.4, we have $\chi_c(z; F) \geqslant 0$ for μ_∞-almost every point z. Thus we can apply Proposition 4.10 to the situation where $X(i) := \mathcal{B}(\mu_i)$ and $\mu_{i,\infty} = \mu_i$ for $i \geqslant 1$. Since the μ_i's are ergodic, the disks D_i in the corresponding conclusion are contained in $\mathcal{B}(\mu_i)$ modulo Lebesgue null subsets and hence mutually disjoint. But this is impossible because the radii of the disks D_i are positive and independent of i. □

Proof of Proposition 4.9. Let $X = M \setminus (\mathcal{B}^-(F) \cup \mathcal{B}^+(F))$. For the proof of the proposition, it is enough to show the claim in the case when $\mathbf{m}(X) > 0$ and $\nu = \mathbf{m}_X$. Let ν_∞ be a weak limit point of the sequence $n^{-1} \sum_{j=0}^{n-1} \nu \circ F^{-j}$. From Lemma 4.6, $\chi_c(z; F) \geqslant 0$ for ν_∞-almost every $z \in M$. Thus we have only to prove $\chi_c(\nu_\infty; F) \leqslant 0$. Suppose that we have $\chi_c(\nu_\infty; F) > 0$. Then we can apply Proposition 4.10 to the situation where $X(i) := X$ for all $i \geqslant 1$. Let $\nu_{i,\infty} \ll \nu_\infty$ and D_i be those in the corresponding conclusion, which we can assume to be independent of i. We should have

$$\nu(\mathcal{B}(\nu_{i,\infty})) \geqslant \limsup_{n\to\infty} \nu(F^{-n}(D_i)) \geqslant \nu_\infty(D_i) > 0.$$

But this contradicts the definition of X because $\nu_{i,\infty}$ is an ergodic physical measure with positive central Lyapunov exponent. □

We proceed to the proof of Proposition 4.10. For positive numbers χ, ε, k and a positive integer n, we define a closed subset $\Gamma(\chi, \varepsilon, k, n; F)$ as the set of all points $z \in M$ such that, for any $0 \leqslant m < n$ and any $v \in \mathbf{S}^u(F^m(z))$,

(Γ1) $|D^*F^{n-m}(v)| \geqslant \exp(\chi(n-m) - k)$;

(Γ2) $|D^*F(v)| \geqslant \exp(-\varepsilon(n-m) - k)$.

For the points in $\Gamma(\chi, \varepsilon, k, n; F)$, we have the following estimates on distortion:

LEMMA 4.11. *For positive numbers $\chi > 0$, $0 < \varepsilon < \frac{1}{10}\chi$ and $k > 0$, there exists a positive constant $\alpha = \alpha(\chi, \varepsilon, k)$, which depends only on χ, ε and k besides the objects that we fixed at the end of* §3.2, *such that, for any $n > 0$ and $z \in \Gamma(\chi, \varepsilon, k, n; F)$, the restriction of F^n to some neighborhood V of z is a diffeomorphism onto the disk $\mathbf{B}(F^n(z), \alpha)$ and we have*

(1) $\|(DF^n_w)^{-1}\|^{-1} > C_g^{-1} \exp(\chi n - k)$ *for* $w \in V$;

(2) $\big|\log|\det DF^n_w| - \log|\det DF^n_{w'}|\big| < 1$ *for* $w, w' \in V$.

Proof. Fix $v \in \mathbf{S}^u(z)$ and put $\delta(i) = |D^*F^{n-i}(DF^i(v))|^{-1}$ for $0 \leqslant i < n$. Let D_n be the disk in the tangent space $T_{F^n(z)}M$ with center at the origin and radius α. We define the regions $D_i \subset T_{F^i(z)}M$ for $0 \leqslant i < n$ so that $DF(D_i)$ is the $\delta(i)\alpha$-neighborhood of D_{i+1}. Then we have

$$\operatorname{diam} D_i \leqslant \|(DF^{n-i}_{F^i(z)})^{-1}\|\, \alpha + \sum_{j=i}^{n-1} \|(DF^{j+1-i}_{F^i(z)})^{-1}\|\, \delta(j)\alpha$$

for $0\leqslant i<n$. Using the relation (7), we can check that

$$\|(DF^{j+1-i}_{F^i(z)})^{-1}\|\,\delta(j)\leqslant C_g|D^*F(DF^j(v))|^{-1}\delta(i).$$

Thus, from the conditions (Γ1) and (Γ2), we get

$$\operatorname{diam} D_i\leqslant C_g(n-i+1)\exp(\varepsilon(n-i)+k)\,\delta(i)\,\alpha\leqslant C_g(n-i+1)\exp(-(\chi-\varepsilon)(n-i)+2k)\,\alpha.$$

From the condition (Γ2) and the relation (7), we have

$$\|DF^{-1}_{F^i(z)}\|^{-1}\geqslant C_g^{-1}\exp(-\varepsilon(n-i)-k).$$

For $v\in D_i$, we have the estimates

$$\begin{aligned}\|\exp^{-1}_{F^{i+1}(z)}\circ F\circ\exp_{F^i(z)}(v)-DF_{F^i(z)}(v)\|&\leqslant C_g(\operatorname{diam} D_i)^2\\ &\leqslant C_gn^2\exp(-(\chi-2\varepsilon)(n-i)+3k)\,\delta(i)\,\alpha^2\end{aligned}$$

and

$$\begin{aligned}\|D(\exp^{-1}_{F^{i+1}(z)}\circ F\circ\exp_{F^i(z)})_v-DF_{F^i(z)}\|&\leqslant C_g\operatorname{diam} D_i\\ &\leqslant C_gn\exp(-(\chi-\varepsilon)(n-i)+2k)\,\alpha.\end{aligned}$$

Hence, if we take sufficiently small α depending only on χ, ε, k and C_g, the restriction of F to $\exp_{F^i(z)}(D_i)$ is a diffeomorphism onto a neighborhood of the subset $\exp_{F^{i+1}(z)}(D_{i+1})$ for $0\leqslant i<n$. This implies the first claim of the lemma. We can check, by straightforward estimates, that the other claims, (1) and (2), hold if we take sufficiently small α. □

From now to the end of this section, we consider the situation in Proposition 4.10. For each i, we take a subsequence $n(j;i)\to\infty$ $(j\to\infty)$ such that the sequence of measures $n(j;i)^{-1}\sum_{m=0}^{n(j;i)-1}\mathbf{m}_{X(i)}\circ F^{-m}$ converges to $\mu_{i,\infty}$ as $j\to\infty$. The following is the key lemma in the proof of Proposition 4.10:

LEMMA 4.12. *There exist $\chi>0$, $0<\varepsilon<\frac{1}{10}\chi$ and $k>0$ such that*

$$\liminf_{j\to\infty}\frac{1}{n(j;i)}\sum_{m=0}^{n(j;i)-1}\mathbf{m}_{X(i)}(\Gamma(\chi,\varepsilon,k,m;F))>0\quad\textit{for sufficiently large } i. \tag{31}$$

The point of this lemma is that we can take χ, ε and k uniformly for sufficiently large i. Before proving this lemma, we finish the proof of Proposition 4.10 assuming it.

Proof of Proposition 4.10. Let the constants χ, ε and k be those in Lemma 4.12 and $\alpha=\alpha(\chi,\varepsilon,k,F)$ that in Lemma 4.11. We consider a large integer i for which (31) holds. Then we can take a compact subset $K\subset X(i)$ and a point $z_0\in M$ such that

$$\liminf_{j\to\infty}\frac{1}{n(j;i)}\sum_{m=0}^{n(j;i)-1}(\mathbf{m}|_{K\cap\Gamma(\chi,\varepsilon,k,m;F)}\circ F^{-m})\big(\mathbf{B}\big(z_0,\tfrac{1}{2}\alpha\big)\big)>0. \tag{32}$$

Let $\mathcal{D}_m$ be the union of the connected components of $F^{-m}\big(\mathbf{B}\big(z_0,\frac{1}{2}\alpha\big)\big)$ that meet $K\cap\Gamma(\chi,\varepsilon,k,m;F)$. Then, on each of the connected components of $\mathcal{D}_m$, the mapping F^n is a diffeomorphism onto $\mathbf{B}\big(z_0,\frac{1}{2}\alpha\big)$ and satisfies the estimates in Lemma 4.11. Let ν_i be a limit point of the sequence $\{n(j;i)^{-1}\sum_{m=0}^{n(j;i)-1}\mathbf{m}|_{\mathcal{D}_m}\circ F^{-m}\}_{j=1}^{\infty}$. Then we have $\nu_i\leqslant\mathbf{m}(X(i))\,\mu_{i,\infty}$ and $\nu_i\ll\mathbf{m}$, and, further,

$$e^{-1}\frac{\nu_i\big(\mathbf{B}\big(z_0,\frac{1}{2}\alpha\big)\big)}{\mathbf{m}\big(\mathbf{B}\big(z_0,\frac{1}{2}\alpha\big)\big)}\leqslant\frac{d\nu_i}{d\mathbf{m}}\leqslant e\frac{\nu_i\big(\mathbf{B}\big(z_0,\frac{1}{2}\alpha\big)\big)}{\mathbf{m}\big(\mathbf{B}\big(z_0,\frac{1}{2}\alpha\big)\big)}.$$

We can check that ν_i is ergodic and $\chi_c(z;F)>0$ for ν_i-almost every point z. (See the remark below.) Hence there is an ergodic component $\nu_{i,\infty}$ of $\mu_{i,\infty}$ such that $\nu_i\ll\nu_{i,\infty}\ll\mu_{i,\infty}$. The measure $\nu_{i,\infty}$ and the disk $D_i=\mathbf{B}\big(z_0,\frac{1}{2}\alpha\big)$ satisfy the conditions in Proposition 4.10. □

Remark. Actually, it is not completely simple to prove that the measure ν_i in the proof above is ergodic and that $\chi_c(z;F)>0$ for ν_i-almost every point z. But there are a few standard ways for it. For example, we can argue as follows: Consider the inverse limit space of F,

$$\widetilde{M}_F=\{\{z_j\}_{j=-\infty}^{0}\mid z_j\in M\text{ and }F(z_j)=z_{j+1}\},$$

and the projection $\pi:\widetilde{M}_F\to M$ defined by $\pi(\{z_j\}_{j=-\infty}^{0})=z_0$. Let $\tilde{\mu}_{i,\infty}$ be the natural extension of $\mu_{i,\infty}$. We can check that the part $\tilde{\nu}_i$ of $\tilde{\mu}_{i,\infty}$ that corresponds to ν_i is supported on a union of local unstable manifolds, each of which is projected onto the disk $\mathbf{B}\big(z_0,\frac{1}{2}\alpha\big)$ by π. Further, the conditional measures on those local unstable manifolds given by $\tilde{\nu}_i$ are absolutely continuous with respect to the smooth measures on them. For any continuous function φ on M, the backward time average of $\varphi\circ\pi$ is constant on each of the local unstable manifolds. From the ergodic theorem, the forward time average coincides with the backward time average almost everywhere with respect to $\tilde{\nu}_i\ll\tilde{\mu}_{i,\infty}$, and is the pullback of a function on M by π. Thus it must be constant $\tilde{\nu}_i$-almost everywhere. This implies that ν_i is ergodic. The positivity of the central Lyapunov exponent is obtained by considering Lyapunov exponents with respect to the backward iteration.

In the remaining part of this subsection, we prove Lemma 4.12. To begin with, we fix several constants: Fix $\chi_0>0$ and $0<s_0<1$ such that

$$\mu_\infty(\{z\in M \mid \chi_c(z)>\chi_0\})>s_0. \tag{33}$$

Also fix a positive number ε_0 such that $0<\varepsilon_0<10^{-4}s_0\chi_0$. Recall that we are considering a mapping $F\in\mathcal{U}$ that satisfies the no flat contact condition. From Lemma 3.15, we can fix a large positive constant $h_0>\chi_0$ such that

$$\int \min\{0, L(z;F)+h_0\}\, d(\mu\circ F^{-n})(z) > -\frac{1}{10}s_0\varepsilon_0$$

for any measure μ in $\mathcal{AM}([1,\infty))$ and $n\geqslant n_0(F)$, where $L(z;F)$ is the function defined by (17) and $n_0(F)$ is the constant in the definition of the no flat contact condition. From (33) and the assumption that $\chi_c(z;F)\geqslant 0$ for μ_∞-almost every z, we can fix a constant $k_0>h_0$ such that

$$\mu_\infty(\{z\in M \mid |D^*F^n(v)|\geqslant \exp(\chi_0 n-k_0) \text{ for all } v\in \mathbf{S}^u(z) \text{ and } n\geqslant 0\})>s_0,$$
$$\mu_\infty(\{z\in M \mid |D^*F^n(v)|\geqslant \exp(-\varepsilon_0 n-k_0) \text{ for all } v\in \mathbf{S}^u(z) \text{ and } n\geqslant 0\})>1-\frac{s_0\varepsilon_0}{10h_0}.$$

Finally we fix a positive integer m_0 that satisfies $\varepsilon_0 m_0>10k_0$.

Next we introduce the following subsets of M:

$$A=\{z\in M \mid |D^*F^m(v)|>\exp(\chi_0 m-2k_0) \text{ for all } v\in \mathbf{S}^u(z) \text{ and } 0\leqslant m\leqslant m_0\},$$
$$B=\{z\in M \mid |D^*F^m(v)|>\exp(-\varepsilon_0 m-2k_0) \text{ for all } v\in \mathbf{S}^u(z) \text{ and } 0\leqslant m\leqslant m_0\}\supset A,$$
$$C=M\setminus B,$$
$$D=\{z\in C \mid L(z;F)\leqslant -h_0\}\subset C.$$

Note that A and B are open subsets. From the assumption that the sequence $\mu_{i,\infty}$ converges to μ_∞ as $i\to\infty$, we have

$$\liminf_{j\to\infty}\frac{1}{n(j;i)}\sum_{m=0}^{n(j;i)-1}\mathbf{m}_{X(i)}(F^{-m}(A))>s_0, \tag{34}$$

$$\liminf_{j\to\infty}\frac{1}{n(j;i)}\sum_{m=0}^{n(j;i)-1}\mathbf{m}_{X(i)}(F^{-m}(B))>1-\frac{s_0\varepsilon_0}{10h_0} \tag{35}$$

for sufficiently large i.

We fix a large integer i for which (34) and (35) hold. Using Lemma 3.12, we can find a small number $b_0>0$ and a probability measure μ_0 in $\mathcal{AM}([b_0,\infty))$ such that

$$|\mathbf{m}_{X(i)}-\mu_0|<\tfrac{1}{10}s_0, \tag{36}$$

$$\liminf_{j\to\infty}\frac{1}{n(j;i)}\sum_{m=0}^{n(j;i)-1}\mu_0(F^{-m}(A))>s_0, \tag{37}$$

$$\liminf_{j\to\infty}\frac{1}{n(j;i)}\sum_{m=0}^{n(j;i)-1}\mu_0(F^{-m}(B))>1-\frac{s_0\varepsilon_0}{10h_0}. \tag{38}$$

By modifying the measure μ_0 slightly if necessary, we can assume that

$$\sum_{m=0}^{n_0(F)}\int\min\{0,L(F^m(z);F)+h_0\}\,d\mu_0>-\infty$$

in addition. Then, from Corollary 3.8 and the choice of h_0, we also have

$$\liminf_{j\to\infty}\frac{1}{n(j;i)}\sum_{m=0}^{n(j;i)-1}\int\min\{0,L(F^m(z);F)+h_0\}\,d\mu_0\geqslant-\frac{s_0\varepsilon_0}{10}. \tag{39}$$

For $z\in M$ and integers $m<n$, let $A_z(m,n)$, $B_z(m,n)$, $C_z(m,n)$ and $D_z(m,n)$, be the set of integers $m\leqslant q<n$ for which $F^q(z)$ belongs to A, B, C and D, respectively. Then we have the following result:

LEMMA 4.13. *A point $z\in M$ belongs to $\Gamma\left(\frac{1}{40}s_0\chi_0,4\varepsilon_0,6k_0,n;F\right)$ for $n>0$ if*
(A) *$\#A_z(m,n)\geqslant\frac{1}{10}s_0(n-m)$ for any $0\leqslant m<n$;*
(C) *$\#C_z(m,n)\leqslant\varepsilon_0(n-m)/h_0$ for any $0\leqslant m<n$;*
(D) *$\sum_{q\in D_z(m,n)}\min\{0,L(F^q(z);F)+h_0\}\geqslant-\varepsilon_0(n-m)$ for any $0\leqslant m<n$.*

Proof. Consider a point $z\in M$ and an integer n that satisfy the conditions (A), (C) and (D). Let $0\leqslant m<n$ and $I=\{m,m+1,...,n-1\}$. We call a set of m_0 consecutive integers $\{q,q+1,...,q+m_0-1\}$ an A-interval (resp. a B-interval) if its smallest element q belongs to $A_z(m,n)$ (resp. $B_z(m,n)$). If $\{q,q+1,...,q+m_0-1\}$ is an A-interval, we have

$$\sum_{j=0}^{m_0-1}\log|D^*F(DF^j(v))|\geqslant\chi_0m_0-2k_0>(\chi_0-\varepsilon_0)m_0+2k_0 \tag{40}$$

for $v\in\mathbf{S}^u(F^q(z))$, where the second inequality follows from the choice of m_0. Similarly, if $\{q,q+1,...,q+m_0-1\}$ is a B-interval, we have

$$\sum_{j=0}^{m_0-1}\log|D^*F(DF^j(v))|\geqslant-\varepsilon_0m_0-2k_0>-2\varepsilon_0m_0 \tag{41}$$

for $v \in \mathbf{S}^u(F^q(z))$.

Take mutually disjoint A-intervals that cover $A_z(m,n)$, and let I_A be the union of them. Then take mutually disjoint B-intervals that cover $B_z(m,n)\setminus I_A$, and let I_B be the union of them. We can take the B-intervals in I_B so that their smallest elements are *not* contained in I_A. Note that I_A and I_B are not necessarily contained in I.

Consider an arbitrary vector $v \in \mathbf{S}^u(F^m(z))$. Then $DF^{q-m}(v)$ belongs to $\mathbf{S}^u(F^q(z))$ for $q \geqslant m$. From (40) and the fact that all the A-intervals in I_A but one is contained in I, we have

$$\sum_{q \in I_A \cap I} \log |D^*F(DF^{q-m}(v))| \geqslant (\chi_0 - \varepsilon_0)\#(I_A \cap I) + 2k_0(\#I_A/m_0 - 1) - 2k_0.$$

Each A-interval in I_A meets at most one B-interval in I_B. Thus the number of B-intervals in I_B whose intersection with $I\setminus I_A$ has cardinality less than m_0 is at most $\#I_A/m_0+1$. From this and (41), we obtain

$$\sum_{q \in I_B \cap (I \setminus I_A)} \log |D^*F(DF^{q-m}(v))| \geqslant -2\varepsilon_0 \#(I_B \cap (I\setminus I_A)) - 2k_0(\#I_A/m_0+1).$$

Since the complement of $I_A \cup I_B$ in I is contained in $C_z(m,n)$, the condition (Γ1) in the definition of the set $\Gamma\left(\frac{1}{40}s_0\chi_0, 4\varepsilon_0, 6k_0, n; F\right)$ follows from the two inequalities above, the assumptions (A), (C) and (D), and the choice of ε_0. If m belongs to $B_z(m,n)$, the condition (Γ2) obviously holds. Otherwise, the condition (Γ2) follows from (D) because we have $\varepsilon_0(n-m)/h_0 \geqslant \#C_z(m,n) \geqslant 1$ in that case from (C). □

In order to prove Lemma 4.12, we see how often the assumptions (A), (C) and (D) in the lemma above hold. For this purpose, we prepare the following elementary lemma, which we shall use again in §6:

LEMMA 4.14. *Let μ be a measure on a measurable space X and ψ_m, $m=0,1,\dots$, be a sequence of non-negative-valued integrable functions on X. For a positive number $\alpha>0$ and an integer $p\geqslant 0$, let $Y_p(\alpha)$ be the set of points $y \in X$ such that*

$$\sum_{l=q}^{p-1} \psi_l(y) \geqslant \alpha(p-q) \quad \text{for some } 0 \leqslant q < p.$$

(So $Y_0(\alpha)=\varnothing$.) Then, for any $n>0$,

$$\sum_{m=0}^{n-1} \mu(Y_m(\alpha)) \leqslant \sum_{m=0}^{n} \mu(Y_m(\alpha)) \leqslant \frac{1}{\alpha} \sum_{m=0}^{n-1} \int \psi_m \, d\mu.$$

Proof. For each point $z \in M$, we define integers

$$n = q_0(z) \geqslant p_1(z) > q_1(z) \geqslant p_2(z) > q_2(z) \geqslant \ldots \geqslant p_{j(z)} > q_{j(z)} \geqslant 0$$

in the following inductive manner: Suppose that $q_j(z)$ has been defined. If there exist integers $p \leqslant q_j(z)$ such that $z \in Y_p(\alpha)$, let $p_{j+1}(z)$ be the maximum of these integers and $q_{j+1}(z)$ the smallest integer $q < p_{j+1}(z)$ such that

$$\sum_{l=q}^{p_{j+1}(z)-1} \psi_l(z) \geqslant \alpha(p_{j+1}(z)-q). \tag{42}$$

Otherwise we put $j(z)=j$ and finish the definition. Consider the subsets

$$Z_m = \{z \in M \mid q_j(z) \leqslant m < p_j(z) \text{ for some } 1 \leqslant j \leqslant j(z)\}$$

for $0 \leqslant m < n$. Then we have $Y_{m+1}(\alpha) \subset Z_m$. From (42), we obtain

$$\sum_{m=0}^{n-1} \int \psi_m \, d\mu \geqslant \alpha \sum_{m=0}^{n-1} \mu(Z_m) \geqslant \alpha \sum_{m=1}^{n} \mu(Y_m(\alpha)) = \alpha \sum_{m=0}^{n} \mu(Y_m(\alpha)). \qquad \square$$

Now we can complete the proof of Lemma 4.12.

Proof of Lemma 4.12. For $n \geqslant 0$, let $\tilde{A}_n$, $\widetilde{C}_n$ and $\widetilde{D}_n$ be the set of points $z \in M$ for which the condition (A), (C) and (D) does *not* hold, respectively. First, apply Lemma 4.14 to the case where $\alpha = 1 - \frac{1}{10}s_0$, $n = n(j;i)$ and ψ_m is the indicator function of the complement of $F^{-m}(A)$. Then, from (37), we obtain

$$\frac{1}{n(j;i)} \sum_{m=0}^{n(j;i)-1} \mu_0(\tilde{A}_m(z)) \leqslant \frac{1}{1-\frac{1}{10}s_0} \frac{1}{n(j;i)} \sum_{m=0}^{n(j;i)-1} \mu_0(M \setminus F^{-m}(A)) \leqslant \frac{1-s_0}{1-\frac{1}{10}s_0} \leqslant 1 - \frac{9}{10}s_0$$

for sufficiently large j. Second, apply Lemma 4.14 to the case where $\alpha = \varepsilon_0/h_0$, $n = n(j;i)$ and ψ_m is the indicator function of the set $F^{-m}(C) = M \setminus F^{-m}(B)$. Then, from (38), we obtain

$$\frac{1}{n(j;i)} \sum_{m=0}^{n(j;i)-1} \mu_0(\widetilde{C}_m(z)) \leqslant \frac{h_0}{\varepsilon_0} \frac{1}{n(j;i)} \sum_{m=0}^{n(j;i)-1} \mu_0(F^{-m}(C)) \leqslant \frac{1}{10}s_0$$

for sufficiently large j. Third, apply Lemma 4.14 to the case where $\alpha = \varepsilon_0$, $n = n(j;i)$ and $\psi_m(z) = -\min\{0, L(F^m(z);F) + h_0\}$. Then, from (39), we obtain

$$\frac{1}{n(j;i)} \sum_{m=0}^{n(j;i)-1} \mu_0(\widetilde{D}_m(z)) \leqslant -\frac{1}{\varepsilon_0} \frac{1}{n(j;i)} \sum_{m=0}^{n(j;i)-1} \int \min\{0, L(F^m(z);F) + h_0\} \, d\mu_0(z) \leqslant \frac{1}{10}s_0$$

for sufficiently large j. From the three inequalities above and (36), we conclude that

$$\frac{1}{n(j;i)} \sum_{m=0}^{n(j;i)-1} \mathbf{m}_{X(i)}(\tilde{A}_m \cup \widetilde{C}_m \cup \widetilde{D}_m) \leqslant 1 - \frac{6}{10}s_0$$

for sufficiently large j. Since the complement of $\tilde{A}_m \cup \widetilde{C}_m \cup \widetilde{D}_m$ is contained in the subset $\Gamma\big(\frac{1}{40}s_0\chi_0, 4\varepsilon_0, 6k_0, m; F\big)$ from Lemma 4.13, this implies Lemma 4.12. $\square$

5. Some estimates on distortion

In this section, we give some basic estimates on distortion of the iterates of mappings in $\mathcal{U}$. The estimates are straightforward and may look rather tedious. But we need to check that some constants in the estimates can be taken uniformly for the mappings in $\mathcal{U}$. This is important especially in our argument in §7, where we consider perturbations of mappings in $\mathcal{U}$.

Let $\chi=\{\chi_c^-,\chi_c^+,\chi_u^-,\chi_u^+\}$ be a quadruple satisfying (18), (28) and $\chi_c^-+\chi_u^->0$, and let $\varepsilon>0$ be a small positive number satisfying

$$\varepsilon<10^{-3}\min\{\chi_c^-+\chi_u^-,|\chi_c^-|,\chi_u^+-\chi_u^-,\lambda_g\}. \tag{43}$$

In the argument below, we will take several constants that depend only on χ and ε besides the integer r and the objects that we have already fixed in §3.2. In order to distinguish such constants, we will use symbols with subscript ε for them. Also we will use a generic symbol C_ε for large positive constants of this kind. The usage of this notation is the same as the one introduced in §3.3. The following lemma is the main ingredient of this section:

LEMMA 5.1. *There exist positive constants* $0<\varrho_\varepsilon<1$, $\varkappa_\varepsilon>1$ *and* $\varkappa_g>1$ *such that the following claim holds for any* $F\in\mathcal{U}$, $k>0$, $n\geqslant1$, $z_0\in\Lambda(\chi,\varepsilon,k,n;F)$ *and* $0<\varrho\leqslant\varrho_0$, *where*

$$\varrho_0:=\varrho_\varepsilon e^{-4\varepsilon n-2k}\min_{0\leqslant i\leqslant j\leqslant n}\min_{v\in\mathbf{S}^u(F^i(z_0))}|D^*F^{j-i}(v)|\geqslant\varrho_\varepsilon\exp((\chi_c^--5\varepsilon)n-3k). \tag{44}$$

For every mapping $G\in C^r(M,M)$ *that satisfies* $d_{C^1}(F,G)\leqslant\varrho$, *we can take a point* $z(G)$ *and a neighborhood* $V_\varrho(G)\ni z(G)$ *in a unique manner so that*

(i) $z(G)$ *depends on* G *continuously and* $z(F)=z_0$;

(ii) $G^n(z(G))\equiv F^n(z_0)$;

(iii) *the restriction of* G^n *to* $V_\varrho(G)$ *is a diffeomorphism onto* $\mathbf{B}(F^n(z_0),\varrho)$.

Further it holds that

(iv) $\operatorname{diam}V_\varrho(G)<\varkappa_g\varrho\exp(-\chi_c^-n+k)$;

(v) $\mathbf{B}(z(G),\varkappa_g^{-1}\varrho\exp(-\chi_u^+n-k))\subset V_\varrho(G)$;

(vi) $V_\varrho(G)\subset\Lambda(\chi,\varepsilon,k+1,n;F)$;

(vii) $\angle(DG^n(\mathbf{E}^u(w)),DF^n(\mathbf{E}^u(z_0)))\leqslant\varkappa_\varepsilon e^{2k}\varrho$ *for any point* $w\in V_\varrho(G)$;

(viii) *any admissible curve in* $\mathbf{B}(z_0,\varkappa_g^{-1})$ *meets* $V_\varrho(F)$ *in a single curve.*

Proof. First of all, notice that the inequality in (44) follows from the assumption $z_0\in\Lambda(\chi,\varepsilon,k,n;F)$. We will give the conditions on the choice of the constants ϱ_ε, $\varkappa_\varepsilon$ and $\varkappa_g$ in the course of the argument below. For $0\leqslant i\leqslant n$, we put $\zeta(i)=F^i(z_0)$ and

$$\delta(i)=\frac{\varrho\exp(\varepsilon(n-i)+k)}{\min_{i\leqslant l\leqslant n}\min_{v\in\mathbf{S}^u(\zeta(i))}|D^*F^{l-i}(v)|}.$$

Then we have

$$\varrho < \varrho \exp(\varepsilon(n-i)+k) \leqslant \delta(i) \leqslant \varrho_\varepsilon \exp(-3\varepsilon n - k) \quad \text{for } 0 \leqslant i \leqslant n. \tag{45}$$

Using the relation (7), we can see that

$$\begin{aligned}\frac{\delta(j)}{\delta(i)} &\leqslant \exp(-\varepsilon(j-i)) \frac{\min_{j\leqslant l\leqslant n} \min_{v\in \mathbf{S}^u(\zeta(i))} |D^*F^{l-i}(v)|}{\min_{j\leqslant l\leqslant n} \min_{v\in \mathbf{S}^u(\zeta(j))} |D^*F^{l-j}(v)|} \\ &\leqslant C_g \exp(-\varepsilon(j-i)) \|(DF^{j-i}_{\zeta(i)})^{-1}\|^{-1}\end{aligned} \tag{46}$$

for $0\leqslant i\leqslant j\leqslant n$, and

$$\begin{aligned}\frac{\delta(i+1)}{\delta(i)} &= \exp(-\varepsilon) \frac{\min\{1, \min_{i+1\leqslant l\leqslant n} \min_{v\in \mathbf{S}^u(\zeta(i))} |D^*F^{l-i}(v)|\}}{\min_{i+1\leqslant l\leqslant n} \min_{v\in \mathbf{S}^u(\zeta(i+1))} |D^*F^{l-i-1}(v)|} \\ &\geqslant C_g^{-1} \exp(-\varepsilon) \|(DF_{\zeta(i)})^{-1}\|^{-1}\end{aligned} \tag{47}$$

for $0\leqslant i\leqslant n$.

We put $D_n=\mathbf{B}(0,\varrho)\subset T_{\zeta(n)}M$ and define the region $D_i\subset T_{\zeta(i)}M$ for $0\leqslant i<n$ inductively so that $DF_{\zeta(i)}(D_i)$ is the $2\delta(i+1)$-neighborhood of $D_{i+1}\subset T_{\zeta(i+1)}M$. Put $B_i=\exp_{\zeta(i)}(D_i)$. Then

$$\begin{aligned}\operatorname{diam} B_i = \operatorname{diam} D_i &\leqslant 2\varrho \|(DF^{n-i}_{\zeta(i)})^{-1}\| + \sum_{j=i+1}^{n} 4\delta(j) \|(DF^{j-i}_{\zeta(i)})^{-1}\| \\ &< C_\varepsilon \delta(i) \leqslant C_\varepsilon \varrho_\varepsilon \exp(-3\varepsilon n - k)\end{aligned} \tag{48}$$

for $0\leqslant i\leqslant n$, where the second inequality follows from (46) and the third from (45). Since $\zeta(0)=z_0\in\Lambda(\chi,\varepsilon,k,n;F)$, we have

$$\|(DF_{\zeta(i)})^{-1}\|^{-1} \geqslant C_g^{-1} \exp(-\varepsilon(n-i)-k) \quad \text{for } 0\leqslant i\leqslant n \tag{49}$$

by (7). Therefore, if we take the constant ϱ_ε sufficiently small, we can obtain

$$\|DG_w - DF_{\zeta(i)}\| \leqslant d_{C^1}(F,G) + C_g \operatorname{diam} B_i < \|(DF_{\zeta(i)})^{-1}\|^{-1}$$

and

$$d(G(w), \exp_{\zeta(i+1)} \circ DF_{\zeta(i)} \circ \exp^{-1}_{\zeta(i)}(w)) \leqslant d_{C^1}(F,G) + C_g (\operatorname{diam} B_i)^2 < 2\delta(i+1)$$

for $0\leqslant i<n$, $w\in\mathbf{B}(\zeta(i),\operatorname{diam} B_i)$ and any mapping $G\in C^r(M,M)$ satisfying $d_{C^1}(F,G)\leqslant \varrho\leqslant\varrho_0$, where we have used the relation

$$(\operatorname{diam} B_i)^2 \leqslant C_\varepsilon \delta(i)^2 \leqslant C_\varepsilon \varrho_\varepsilon \exp(-2\varepsilon n)\delta(i+1),$$

which follows from (45), (47) and (49). These two inequalities imply that the mapping G restricted to $\mathbf{B}(\zeta(i), \operatorname{diam} B_i) \supset B_i$ is a diffeomorphism and maps B_i onto a neighborhood of B_{i+1} for $0 \leqslant i < n$. Put $V_\varrho(G) = \bigcap_{i=0}^n G^{-i}(B_i)$. Then the restriction of G^n to $V_\varrho(G)$ is a diffeomorphism onto $B_n = \mathbf{B}(F^n(z_0), \varrho)$. Let $z(G)$ be the unique point in $V_\varrho(G)$ that G^n brings to $F^n(z_0)$. Clearly $z(G)$ and $V_\varrho(G)$ satisfy the conditions (i), (ii) and (iii).

We show the conditions (iv)–(viii). Using (6) and (7), we can check that (iv) and (v) follow from (vi). We prove (vi) and (vii). Let $G \in \mathcal{U}$ be a mapping that satisfies $d_{C^1}(F, G) \leqslant \varrho \leqslant \varrho_0$ and w a point in $V_\varrho(G)$. We put $w(i) = G^i(w)$ for $0 \leqslant i \leqslant n$. Consider an integer $0 \leqslant i \leqslant n$ and tangent vectors $v \in \mathbf{S}^u(\zeta(i))$ and $u \in \mathbf{S}^u(w(i))$. For $0 \leqslant m \leqslant n-i$, we have

$$\begin{aligned}\angle(DG^m_{w(i)}(u), DF^m_{\zeta(i)}(v)) \leqslant &\angle(DF^m_{\zeta(i)}(u), DF^m_{\zeta(i)}(v)) \\ &+ \sum_{j=1}^{m} \angle(DF^{m-j+1}_{\zeta(i+j-1)}(DG^{j-1}_{w(i)}(u)), DF^{m-j}_{\zeta(i+j)}(DG^j_{w(i)}(u))).\end{aligned}$$

Remark. In the expression above, we identified tangent vectors with their parallel translations and abused the notation slightly. In fact, $DF^{m-j}_{\zeta(i+j)}(DG^j_{w(i)}(u))$ should have been written $DF^{m-j}_{\zeta(i+j)}(\tau(DG^j_{w(i)}(u)))$, where τ is the parallel translation from $w(i+j)$ to $\zeta(i+j)$. We continue to use such identifications below.

Since $w(i+j-1) \in B_{i+j-1}$ and $DG^{j-1}_{w(i)}(u) \in \mathbf{S}^u(w(i+j-1))$, the parallel translation of $DG^{j-1}_{w(i)}(u)$ to $\zeta(i+j-1)$ does not belong to $\mathbf{S}^c(\zeta(i+j-1))$, provided that we take sufficiently small ϱ_ε. Also we have

$$\angle(DF_{\zeta(i+j-1)}(DG^{j-1}_{w(i)}(u)), DG^j_{w(i)}(u)) \leqslant C_g(\operatorname{diam} B_{i+j-1} + d_{C^1}(F, G))$$

for $0 \leqslant j \leqslant n-i$. Using these consequences and (4) in the inequality above, we obtain

$$\begin{aligned}\angle(DG^m_{w(i)}(u), DF^m_{\zeta(i)}(v)) &\leqslant A_g \frac{|D^*F^m(v)|}{D_*F^m(v)} \angle(u, v) \\ &\qquad + C_g \sum_{j=1}^{m} \frac{|D^*F^{m-j}(DF^j(v))|}{D_*F^{m-j}(DF^j(v))} (\operatorname{diam} B_{i+j-1} + \varrho) \\ &\leqslant C_g \exp(-\lambda_g m) \angle(u, v) \\ &\qquad + C_g \sum_{j=1}^{m-1} \exp(-\lambda_g(m-j))(\operatorname{diam} B_{i+j-1} + \varrho).\end{aligned} \tag{50}$$

In order to prove the condition (vii), we consider (50) in the case where $i=0$, $m=n$ and v and u are unit tangent vectors in $\mathbf{E}^u(z_0)$ and $\mathbf{E}^u(w)$, respectively. In this case, we have

$$\begin{aligned}\frac{|D^*F^{n-j}(DF^j(v))|}{D_*F^{n-j}(DF^j(v))}(\operatorname{diam} B_{j-1}+\varrho) &\leqslant \frac{|D^*F^{n-j}(DF^j(v))|}{D_*F^{n-j}(DF^j(v))} C_\varepsilon \delta(j-1)\\ &\leqslant C_\varepsilon \varrho \exp(\varepsilon(n-j)+k)\\ &\qquad \times \max_{j\leqslant l\leqslant n} \frac{|D^*F^{n-l}(DF^l(v))|}{D_*F^{n-l}(DF^l(v))} \frac{|D^*F(DF^{j-1}(v))|^{-1}}{D_*F^{l-j}(DF^j(v))}\\ &\leqslant C_\varepsilon \varrho \exp(-(\lambda_g-2\varepsilon)(n-j)+2k)\end{aligned}$$

for $1\leqslant j\leqslant n$, where we used (45) and (48) in the first inequality, (7) in the second, and the assumption $z_0\in\Lambda(\chi,\varepsilon,k,n;F)$ in the third. Likewise, using the estimate $\angle(v,u)\leqslant C_g d(z_0,w)\leqslant C_g \operatorname{diam} B_0$, we can show that

$$\frac{|D^*F^n(v)|}{D_*F^n(v)}\angle(u,v)\leqslant \frac{|D^*F^n(v)|}{D_*F^n(v)} C_g \operatorname{diam} B_0 \leqslant C_\varepsilon \varrho \exp(-(\lambda_g-2\varepsilon)n+2k).$$

Putting these inequalities in (50), we obtain the condition (vii).

Next we prove the condition (vi). Consider an integer $0\leqslant i\leqslant n$ and a vector $u\in \mathbf{S}^u(w(i))$. Since $w(i)$ belongs to B_i, there is a vector $v\in\mathbf{S}^u(\zeta(i))$ such that $\angle(u,v)< C_g \operatorname{diam} B_i$. From this, (48) and (50), we obtain

$$\begin{aligned}|D^*G(DG^l_{w(i)}(v))-D^*F(DF^l_{\zeta(i)}(u))| &\leqslant C_g(|\det DG_{w(i+l)}-\det DF_{\zeta(i+l)}|\\ &\qquad +|D_*G(DG^l_{w(i)}(v))-D_*F(DF^l_{\zeta(i)}(u))|)\\ &\leqslant C_g(d_{C^1}(F,G)+\operatorname{diam} B_{i+l}\\ &\qquad +\angle(DG^l_{w(i)}(v),DF^l_{\zeta(i)}(u)))\\ &\leqslant C_g \varrho_\varepsilon \exp(-3\varepsilon n-k)\end{aligned}$$

for $0\leqslant l\leqslant n-i-1$. Thus, using (49), we can obtain

$$\log\left|\frac{D^*G^{j-i}(v)}{D^*F^{j-i}(u)}\right| < \sum_{l=0}^{j-i-1}\log\left|\frac{D^*G(DG^l_{w(i)}(v))}{D^*F(DF^l_{\zeta(i)}(u))}\right| < 1 \quad \text{for } 0\leqslant i\leqslant j\leqslant n, \tag{51}$$

provided that we take the constant ϱ_ε sufficiently small. Likewise, we can get

$$\log\left|\frac{D_*G^{j-i}(v)}{D_*F^{j-i}(u)}\right| < 1 \quad \text{for } 0\leqslant i\leqslant j\leqslant n.$$

The condition (vi) follows from these two inequalities and the assumption that z_0 belongs to $\Lambda(\chi,\varepsilon,n,k;F)$.

Finally we check the condition (viii). Let γ be an admissible curve in $\mathbf{B}(z_0, \varkappa_g^{-1})$. From the argument in §3.4, the curvature of $F_*^i\gamma$ for $0 \leqslant i \leqslant n$ is bounded by some constant C_g, even though $F_*^i\gamma$ for $0 \leqslant i \leqslant n_g$ may not be admissible. Thus we can take the constant $\varkappa_g$ so large that the following holds: the intersection of any arc $\widetilde{\gamma}$ in $F_*^i\gamma$ with length less than $4\Lambda_g \varkappa_g^{-1}$ with any ball with diameter not larger than $2\varkappa_g^{-1}$ is a single subarc of $\widetilde{\gamma}$ with length less than $4\varkappa_g^{-1}$. The diameter of B_i is bounded by $2\varkappa_g^{-1}$ provided that we take the constant ϱ_ε sufficiently small. Thus, by induction on $0 \leqslant j \leqslant n$, we can check that $\gamma_j := \gamma \cap (\bigcap_{l=0}^{j} F^{-l}(\mathbf{B}(\zeta(l), \operatorname{diam} B_l)))$ consists of a single arc. We obtain the condition (viii) as the case $j = n$. □

Note that the claim of Lemma 5.1 remains true even if we get the constant ϱ_ε smaller and $\varkappa_\varepsilon$ and $\varkappa_g$ larger. By letting the constant ϱ_ε be smaller and $\varkappa_\varepsilon$ larger if necessary, we can show the following claim in addition:

ADDENDUM TO LEMMA 5.1. *Suppose that $F \in \mathcal{U}$, $n \geqslant 1$ and $k > 0$. Then there exists a neighborhood $W(z)$ for each point $z \in \Lambda(\chi, \varepsilon, k, n; F)$ such that*

(ix) *the restriction of F^n to $W(z)$ is a diffeomorphism onto the image. Further, if $W(z) \cap W(w) \neq \varnothing$ for some $w \in \Lambda(\chi, \varepsilon, k, n; F)$, then F^n is injective on the union $W(z) \cup W(w)$.*

(x) $\mathbf{m}(W(z)) > \varkappa_\varepsilon^{-1} \exp(-(\chi_u^+ + \max\{\chi_c^+, 0\} + 7\varepsilon)n - 6k)$.

Proof. We consider a point $z_0 \in \Lambda(\chi, \varepsilon, k, n; F)$ and continue to use the notation in Lemma 5.1 and its proof. Let γ be the curve in $V_{\varrho_0}(F)$ that F^n maps onto the segment $\{\zeta(n) + t\mathbf{e}^c(\zeta(n)) \mid |t| < \varrho_0\} \subset \mathbf{B}(\zeta(n), \varrho_0)$, where $\mathbf{e}^c(\cdot)$ is a unit vector in $\mathbf{E}^c(\cdot)$. From backward invariance of the central cones $\mathbf{S}^c(\cdot)$, the tangent vectors of γ is contained in the central cones, provided that we take a sufficiently small ϱ_ε. From (51) and (7), the length of $F_*^i\gamma$ satisfies

$$|F_*^i\gamma| < C_g \varrho_0 \|(DF_{\zeta(i)}^{n-i})^{-1}\| < C_g \varrho_\varepsilon \exp(-4\varepsilon n - 2k)$$

and, for the case $i = 0$,

$$\begin{aligned} |\gamma| &> C_g^{-1} \varrho_0 \|(DF_{\zeta(0)}^n)^{-1}\| > C_g^{-1} \min_{0 \leqslant i \leqslant j \leqslant n} \varrho_\varepsilon e^{-4\varepsilon n - 2k} \|(DF_{\zeta(j)}^{n-j})^{-1}\| \cdot \|(DF_{\zeta(0)}^i)^{-1}\| \\ &\geqslant C_g^{-1} \varrho_\varepsilon \exp(-\max\{\chi_c^+, 0\} n - 5\varepsilon n - 4k). \end{aligned}$$

Next consider the family of parallel segments

$$\gamma_y(t) = y + t\mathbf{e}^u(z_0), \quad |t| < \varrho_\varepsilon \exp(-(\chi_u^+ + 2\varepsilon)n - 2k)$$

parameterized by the points $y \in \gamma$, where $\mathbf{e}^u(z_0)$ is a unit vector in $\mathbf{E}^u(z_0)$. We define $W(z_0)$ as the region that this family of segments sweeps. From the estimate on the length of γ above, we can see that $W(z_0)$ satisfies the condition (x), provided that we take a sufficiently large constant $\varkappa_\varepsilon$. Since the mapping F is uniformly expanding in the unstable directions, we can show that

$$|F_*^i\gamma_y| \leqslant C_g\varrho_\varepsilon \exp(-(\chi_u^+ + 2\varepsilon)n - 2k)\, D_*F^i(\mathbf{e}^u(z_0)) < C_g\varrho_\varepsilon \exp(-\varepsilon n - k)$$

for $0 \leqslant i \leqslant n$. Hence the diameter of $F^i(W(z_0))$ is bounded by

$$|F_*^i\gamma| + 2\max_{y\in\gamma} |F_*^i\gamma_y| \leqslant C_g\varrho_\varepsilon \exp(-\varepsilon n - k).$$

If $W(z_0)\cap W(w) \neq \varnothing$ for some point $w \in \Lambda(\chi,\varepsilon,n,k;F)$, the diameter of the image $F^i(W(z_0)\cup W(w))$ is bounded by $4C_g\varrho_\varepsilon \exp(-\varepsilon n - k)$. On the other hand, the distance from $\zeta(i)$ to the critical set $\mathcal{C}(F)$ is not less than $C_g^{-1}\exp(-\varepsilon n - k)$ from (49). Thus, if we take a sufficiently small constant ϱ_ε, the restrictions of F to $F^i(W(z_0)\cup W(w))$ for $0 \leqslant i < n$ are diffeomorphisms, and hence (ix) holds. □

The condition (ix) implies that, if two points z and w in $\Lambda(\chi,\varepsilon,k,n;F)$ satisfy $F^n(z) = F^n(w)$, then the neighborhoods $W(z)$ and $W(w)$ are disjoint. Thus we obtain the following corollary from the condition (x):

COROLLARY 5.2. *For any $F \in \mathcal{U}$, $n \geqslant 1$, $k > 0$ and $\zeta \in M$, we have*

$$\#(\Lambda(\chi,\varepsilon,k,n;F)\cap F^{-n}(\zeta)) \leqslant \varkappa_\varepsilon \exp((\chi_u^+ + \max\{\chi_c^+, 0\} + 7\varepsilon)n + 6k).$$

6. Physical measures with neutral central Lyapunov exponent

In this section, we study physical measures with nearly neutral central Lyapunov exponent. The goal is the proof of Theorem 3.21, which will be carried out in the last three subsections.

6.1. An illustration of the idea of the proof

The argument in this section is based on a new idea that relates the transversality condition on unstable cones to absolute continuity of physical measures with nearly neutral central Lyapunov exponent. In this subsection, we illustrate the idea in a simple example.

As a simplified model of a partially hyperbolic endomorphism, we consider the skew product $F: [0,1)\times\mathbf{R} \to [0,1)\times\mathbf{R}$ defined by

$$F(x,y) = (dx, a_i x + b_i y + c_i) \quad \text{on } [i/d, (i+1)/d)\times\mathbf{R},\ i = 0,1,2,\dots,d-1,$$

where $d \geqslant 2$ is an integer and a_i, b_i and c_i are real numbers. And we assume that

(1) $|b_i| < d$ for $0 \leqslant i < d$, so that F is partially hyperbolic with $\mathbf{E}^c = \langle \partial/\partial y \rangle$;

(2) $|b_i| > d^{-1}$ for $0 \leqslant i < d$, so that F is volume-expanding;

(3) $\sum_{i=0}^{d-1} \log |b_i| < 0$, so that most of the orbits are bounded.

Put $\theta = \max_{1 \leqslant i \leqslant d} |a_i|/(d - |b_i|)$ and $b_{\max} = \max_{1 \leqslant i \leqslant d} |b_i|$. Then F brings a segment with slope less than θ in absolute value to a union of segments with the same property. Assume in addition that

$$|a_i - a_{i'}| > 3\theta b_{\max} \quad \text{for any } i \neq i'. \tag{52}$$

This is a much simplified analogue of the transversality condition on unstable cones. Indeed, if l_σ is a segment in $[i_\sigma/d, (i_\sigma+1)/d) \times \mathbf{R}$ for $\sigma = 1, 2$, and if their slopes are bounded by θ in absolute value, then (52) implies that the difference between the slopes of their images under the mapping F is larger than $\theta b_{\max}/d$, provided $i_1 \neq i_2$.

We prove the existence of an absolutely continuous invariant measure for F with *negative* central Lyapunov exponent. First of all, observe the following fact: if Lebesgue-integrable functions ψ_1 and ψ_2 on $[0,1] \times \mathbf{R}$ take constant values on lines with slopes k_1 and k_2, respectively, or, in other words, satisfy $\psi_i(x, y) = \psi_i(0, y - k_i x)$ for $0 \leqslant x \leqslant 1$ and $y \in \mathbf{R}$, then we have, with $y' = y - k_1 x$,

$$\begin{aligned}(\psi_1, \psi_2)_{L^2} &= \int \psi_1(x, y)\, \psi_2(x, y)\, dx\, dy \\ &= \int \psi_1(0, y')\, \psi_2(0, y' + (k_1 - k_2)x)\, dx\, dy' \\ &\leqslant |k_1 - k_2|^{-1} \|\psi_1\|_{L^1} \|\psi_2\|_{L^1}\end{aligned}$$

provided $k_1 \neq k_2$. Let $\psi(x, y)$ be an L^2-function on $[0,1] \times \mathbf{R}$ and suppose that it is the sum of non-negative functions $\psi_j(y)$, $j = 1, 2, \ldots, m$, that take constant values on lines with slopes k_j with $|k_j| < \theta$, respectively. Let $\mathcal{P}_F$ and $\mathcal{P}_i$, $0 \leqslant i < d$, be the Perron–Frobenius operator associated to F and its restriction to $[i/d, (i+1)/d) \times \mathbf{R}$, respectively, so that $\mathcal{P}_F = \sum_{i=0}^{d-1} \mathcal{P}_i$. By using the transversality condition (52) and the fact that we observed above, we can obtain

$$\|\mathcal{P}_F \psi\|_{L^2}^2 = \sum_{i=0}^{d} \|\mathcal{P}_i \psi\|_{L^2}^2 + \sum_{i \neq i'} (\mathcal{P}_i \psi, \mathcal{P}_{i'} \psi)_{L^2} \leqslant \frac{1}{d \min_i |b_i|} \|\psi\|_{L^2}^2 + \frac{d}{\theta b_{\max}} \|\psi\|_{L^1}^2. \tag{53}$$

Remark. We can regard this inequality as an analogue of the so-called Lasota–Yorke inequality.

Note that the coefficient $1/d \min_i |b_i|$ is smaller than 1 by assumption. The Perron–Frobenius operator $\mathcal{P}_F$ preserves the L^1-norm of non-negative functions and is not dissipative because of the assumption $\sum_{i=0}^{d-1} \log |b_i| < 0$. Since the images $\mathcal{P}_F^n \psi$ again satisfy the condition that we assumed for ψ, we can apply the inequality (53) repeatedly and see that $\mathcal{P}_F^n\psi$, $n=1,2,\ldots$, are uniformly bounded with respect to the L^2-norm. Thus we can find a non-trivial fixed point of $\mathcal{P}_F$ in $L^2([0,1]\times\mathbf{R})$ as an L^2-weak limit point of the sequence $n^{-1}\sum_{m=0}^{n-1}\mathcal{P}_F^m\psi$, $n=1,2,\ldots$. The measure μ having this fixed point as density is an absolutely continuous invariant measure for F, whose central Lyapunov exponent is $d^{-1}\sum_{i=1}^{d}\log|b_i|<0$.

In the argument above, we used the assumption $\sum_{i=1}^{d}\log|b_i|<0$ only to ensure that the Perron–Frobenius operator $\mathcal{P}$ is not dissipative. So, if we consider mappings on compact surfaces, the same argument should be valid in the case where the central Lyapunov exponent is neutral or even slightly positive. This is the key idea that we will develop in the following subsections.

6.2. Semi-norms on the space of measures

For a Borel finite measure μ on M and $0<\delta<1$, we define the function

$$J_\delta\mu\colon \mathbf{T}\longrightarrow\mathbf{R},\quad J_\delta\mu(w):=\frac{\mu(\mathbf{B}(w,\delta))}{\pi\delta^2}=\frac{1}{\pi\delta^2}\int \mathbf{1}_\delta(w,z)\,d\mu(z),$$

where

$$\mathbf{1}_\delta\colon \mathbf{T}\times\mathbf{T}\longrightarrow\mathbf{R},\quad \mathbf{1}_\delta(w,z)=\begin{cases}1, & \text{if } d(w,z)<\delta,\\ 0, & \text{otherwise.}\end{cases}$$

And we put, for Borel finite measures μ and ν on M,

$$(\mu,\nu)_\delta=(J_\delta\mu,J_\delta\nu)_{L^2(\mathbf{m})}\quad\text{and}\quad \|\mu\|_\delta=\sqrt{(\mu,\mu)_\delta}=\|J_\delta\mu\|_{L^2(\mathbf{m})}.$$

Obviously $\|\cdot\|_\delta$ is a semi-norm and satisfies

$$\|\mu\|_\delta\leqslant\frac{|\mu|}{\pi\delta^2}. \tag{54}$$

The semi-norm $\|\mu\|_\delta$ for a measure μ is essentially decreasing with respect to the auxiliary parameter δ. More precisely, we can prove the following lemma:

Lemma 6.1. *There is an absolute constant $C_0>1$ such that*

$$\|\mu\|_\delta\leqslant C_0\|\mu\|_\varrho \tag{55}$$

for any $0<\varrho\leqslant\delta<1$ and any Borel finite measure μ.

Proof. There is an absolute constant C such that, for any $0<\varrho\leqslant\delta<1$, we can cover the disk $\mathbf{B}(0,\delta)$ in $\mathbf{R}^2$ by disks $\mathbf{B}(w_i,\varrho)$, $1\leqslant i\leqslant[C\delta^2/\varrho^2]$, by choosing the points w_i appropriately. Using the Schwarz inequality, we obtain

$$\begin{aligned}\|\mu\|_\delta^2 &= \frac{1}{\pi^2\delta^4}\int \mu(\mathbf{B}(z,\delta))^2\,d\mathbf{m}(z)\\ &\leqslant \frac{1}{\pi^2\delta^4}\int\left(\sum_{i=1}^{[C\delta^2/\varrho^2]}\mu(\mathbf{B}(z+w_i,\varrho))\right)^2 d\mathbf{m}(z)\\ &\leqslant \frac{1}{\pi^2\delta^4}C\frac{\delta^2}{\varrho^2}\sum_{i=1}^{[C\delta^2/\varrho^2]}\int \mu(\mathbf{B}(z+w_i,\varrho))^2\,d\mathbf{m}(z)\\ &\leqslant C^2\|\mu\|_\varrho^2\end{aligned}$$

for any Borel finite measure μ on M. □

We will make use of the following properties of the semi-norm $\|\cdot\|_\delta$:

LEMMA 6.2. *If we have* $\liminf_{\delta\to 0}\|\mu\|_\delta<\infty$ *for a Borel finite measure* μ, *then the measure* μ *is absolutely continuous with respect to the Lebesgue measure* $\mathbf{m}$, *and* $\lim_{\delta\to 0}\|\mu\|_\delta=\|d\mu/d\mathbf{m}\|_{L^2(\mathbf{m})}$.

Proof. The assumption implies that there exists a sequence $\delta(i)\to+0$ such that $J_{\delta(i)}\mu$ is uniformly bounded in $L^2(\mathbf{m})$. Taking a subsequence, we can assume that $J_{\delta(i)}\mu$ converges weakly to some $\psi\in L^2(\mathbf{m})$ as $i\to\infty$. Since

$$(f,\psi)_{L^2(\mathbf{m})}=\lim_{i\to\infty}\int fJ_{\delta(i)}\mu\,d\mathbf{m}=\int f\,d\mu$$

for any continuous function f on M, we have $\mu=\psi\mathbf{m}$. Now the last equality is standard. □

LEMMA 6.3. *If a sequence of Borel finite measures* μ_i, $i\geqslant 1$, *converges weakly to some Borel finite measure* μ_∞, *then we have* $\|\mu_\infty\|_\delta=\lim_{i\to\infty}\|\mu_i\|_\delta$ *for* $\delta>0$.

Proof. We have $\mu_\infty(\partial B(z,\delta))=0$ for Lebesgue almost every point z, because

$$\int\mu_\infty(\partial B(z,\delta))\,d\mathbf{m}(z)=\int_{d(z,w)=\delta}d\mu_\infty(w)\,d\mathbf{m}(z)=\int\mathbf{m}(\partial B(w,\delta))\,d\mu_\infty(w)=0.$$

This implies that $J_\delta\mu_i$ converges to $J_\delta\mu_\infty$ Lebesgue almost everywhere as $i\to\infty$. Since the semi-norms $\|J_\delta\mu_i\|_\delta$, $i\geqslant 1$, are uniformly bounded from (54), the lemma follows from Lebesgue's dominated convergence theorem. □

6.3. Two lemmas on the semi-norm $\|\cdot\|_\delta$

Let $\chi=\{\chi_c^-,\chi_c^+,\chi_u^-,\chi_u^+\}$ be a quadruple satisfying the conditions (18), (28) and (29), and ε a small positive constant satisfying (43). For simplicity, we put

$$\chi_c^\Delta=\chi_c^+-\chi_c^- \quad \text{and} \quad \chi_u^\Delta=\chi_u^+-\chi_u^-.$$

Let F be a mapping in $\mathcal{U}$, k a positive number, n a positive integer and μ a Borel finite measure on M that is supported on the subset $\Lambda(\chi,\varepsilon,k,n;F)$. The aim of this subsection is to give two lemmas that estimate $\|\mu\circ F^{-n}\|_\delta$. Below we shall use the notation in §5.

Suppose that the measure μ is absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$ and that the density $d\mu/d\mathbf{m}$ is square integrable. Then we have

$$\left\|\frac{d(\mu\circ F^{-n})}{d\mathbf{m}}\right\|_{L^2(\mathbf{m})}^2 \leqslant m\exp(-(\chi_c^-+\chi_u^-)n+2k)\left\|\frac{d\mu}{d\mathbf{m}}\right\|_{L^2(\mathbf{m})}^2,$$

where $m=\max\{\#(F^{-n}(w)\cap\Lambda(\chi,\varepsilon,k,n;F))\,|\,w\in M\}$, because

$$|\det DF^n|\geqslant\exp((\chi_c^-+\chi_u^-)n-2k) \quad \text{on } \Lambda(\chi,\varepsilon,k,n;F).$$

The following lemma is a counterpart of this simple fact for the semi-norm $\|\cdot\|_\varrho$. Recall the constants $0<\varrho_\varepsilon<1$ and $\varkappa_\varepsilon,\varkappa_g>1$ in Lemma 5.1.

LEMMA 6.4. *Let ϱ be a positive number satisfying*

$$0<\varrho<\varrho_\varepsilon\frac{\exp((\chi_c^--5\varepsilon)n-3(k+1))}{10\varkappa_g^2}$$

and put

$$\delta=10\varkappa_g\varrho\exp(-\chi_c^-n+k+1).$$

Suppose that a measure μ in $\mathcal{AM}([\delta,\infty))$ is supported on a Borel subset X in $\Lambda(\chi,\varepsilon,k,n;F)$. Then we have

$$\|\mu\circ F^{-n}\|_\varrho^2\leqslant I_g m\exp((-\chi_c^--\chi_u^-+\chi_c^\Delta+\chi_u^\Delta)n+6k)\|\mu\|_\delta^2 \tag{56}$$

for some constant $I_g>0$, where $m=\max\{\#(F^{-n}(w)\cap\mathbf{B}(X,\delta))\,|\,w\in M\}$.

Remark. The point of the lemma above is that the auxiliary parameter of the semi-norm on the right-hand side of (56), that is, δ, is larger than that on the left-hand side, that is, ϱ. If the auxiliary parameter on the right-hand side were allowed to be much smaller than that on the left-hand side, the inequality (56) would hold without the assumption that μ has an admissible lift.

Proof. For each point $y \in \Lambda(\chi, \varepsilon, k+1, n; F)$, there is a unique neighborhood $V(y)$ such that F^n restricted to $V(y)$ is a diffeomorphism onto the disk $\mathbf{B}(F^n(y), \varrho)$, according to Lemma 5.1. Note that the diameter of $V(y)$ is smaller than $\frac{1}{10}\delta$ by Lemma 5.1 (iv) and the definition of δ. Let U be the union of the neighborhoods $V(y)$ for all $y \in X$. Then U is contained in $\mathbf{B}(X, \frac{1}{10}\delta)$ and also in $\Lambda(\chi, \varepsilon, k+1, n; F)$ from Lemma 5.1 (vi) because X is a subset of $\Lambda(\chi, \varepsilon, k, n; F)$. From the definition of U and the assumption that μ is supported on X, it follows that

$$J_\varrho(\mu \circ F^{-n})(w) = \frac{1}{\pi\varrho^2}\mu \circ F^{-n}(\mathbf{B}(w, \varrho)) = \frac{1}{\pi\varrho^2}\sum_{z \in F^{-n}(w) \cap U} \mu(V(z))$$

for $w \in M$. Suppose that we have proved

$$\mu(V(z)) \leqslant C_g \exp(-(\chi_c^- + \chi_u^-)n + 2k)\left(\frac{\varrho}{\delta}\right)^2 \mu(\mathbf{B}(z, \delta)) \tag{57}$$

for any $z \in \Lambda(\chi, \varepsilon, k+1, n; F)$. Then it follows that

$$J_\varrho(\mu \circ F^{-n})(w) \leqslant C_g \exp(-(\chi_c^- + \chi_u^-)n + 2k) \sum_{z \in F^{-n}(w) \cap U} J_\delta \mu(z) \tag{58}$$

for each $w \in M$. As we have

$$|\det DF^n| \leqslant \exp((\chi_c^+ + \chi_u^+)n + 2k + 2) \quad \text{on } U \subset \Lambda(\chi, \varepsilon, k+1, n; F),$$

we can obtain the inequality (56) from (58) by integrating the squares of both sides and using the Schwarz inequality. Therefore, in order to prove the lemma, it is enough to show the inequality (57). Since both sides of (57) are linear with respect to μ, we may assume without loss of generality that μ has an admissible lift that is supported on a single element of the partition $\Xi_{\mathbf{AC}}$ in $\mathbf{AC}([\delta, \infty))$.

Let $\gamma: [0, a] \to M$ be an admissible curve with length $a \geqslant \delta$, and let z be a point in $\Lambda(\chi, \varepsilon, k+1, n; F)$. Consider a connected component I of $\gamma^{-1}(V(z))$, and let J be the connected component of $\gamma^{-1}(\mathbf{B}(z, \delta)) \supset \gamma^{-1}(V(z))$ that contains I. As $\delta < \varkappa_g^{-1}$, Lemma 5.1 (viii) says that the interval I is the unique connected component of $\gamma^{-1}(V(z))$ in J. For the length of I, we have

$$\mathbf{m}_{\mathbf{R}}(I) = |\gamma|_I| \leqslant |F_*^n(\gamma|_I)| \exp(-\chi_u^- n + k + 2) \leqslant C_g \varrho \exp(-\chi_u^- n + k + 2),$$

where the first inequality follows from the fact that $\gamma|_I$ is an admissible curve in $V(z) \subset \Lambda(\chi, \varepsilon, k+2, n; F)$ and the second from the fact that $F_*^n(\gamma|_I)$ is a curve in $F^n(V(z)) = \mathbf{B}(F^n(z), \varrho)$ whose tangent vectors are contained in the unstable cones $\mathbf{S}^u$. For the length of J, we have $\mathbf{m}_{\mathbf{R}}(J)\geqslant\frac{1}{2}\delta$ because the curve $\gamma|_J$ meets $V(z)\subset\mathbf{B}(z,\frac{1}{10}\delta)$ while the length of γ is not less than δ. These estimates hold for each connected component of $\gamma^{-1}(V(z))$. Thus we obtain

$$\frac{\mathbf{m}_{\mathbf{R}}(\gamma^{-1}(V(z)))}{\mathbf{m}_{\mathbf{R}}(\gamma^{-1}(\mathbf{B}(z,\delta)))}<C_g\frac{\varrho\exp(-\chi_u^- n+k)}{\delta}<C_g\frac{\varrho^2}{\delta^2}\exp(-(\chi_c^-+\chi_u^-)n+2k),$$

where we used the definition of δ in the second inequality. From this and the definition of admissible measure, we can conclude (57) for any measure μ that has an admissible lift supported on $\{\gamma\}\times[0,a]$. □

The next lemma is a counterpart of the inequality (53). Recall the definition of $\mathbf{N}(\chi,\varepsilon,k,n;F)$ in §3.7.

Lemma 6.5. *Let ϱ and δ be positive numbers that satisfy*

$$\varrho\exp((-\chi_c^-+\varepsilon)n)\leqslant\delta\leqslant\exp((\chi_c^--2\chi_u^+-3\varepsilon)n).$$

Suppose that a measure μ in $\mathcal{AM}([\delta,\infty))$ is supported on $\Lambda(\chi,\varepsilon,k,n;F)$. Then we have

$$\|\mu\circ F^{-n}\|_\varrho^2\leqslant\frac{\mathbf{N}(\chi,\varepsilon,k+1,n;F)\,\|\mu\|_\varrho^2}{\exp((\chi_c^-+\chi_u^--\chi_c^\Delta-\chi_u^\Delta-2\varepsilon)n)}+\frac{\exp((-2\chi_c^++2\varepsilon)n)}{\delta^2}|\mu|^2,$$

provided that n is larger than some integer $n_*=n_*(\chi,\varepsilon,k)$ which depends only on χ, ε and k besides the objects that we have fixed at the end of §3.2.

Proof. In the course of the proof below, we will give some conditions on the choice of $n_*=n_*(\chi,\varepsilon,k)$. First, we require that n_* is so large that we have

$$\exp((\chi_c^--\chi_u^+-\varepsilon)n_*)<\varrho_\varepsilon\frac{\exp((\chi_c^--5\varepsilon)n_*-3(k+1))}{10\varkappa_g^2}.$$

Consider an integer $n\geqslant n_*$ and put $\varrho_1:=\exp((\chi_c^--\chi_u^+-\varepsilon)n)$. Let $\mathbf{L}(\varrho_1)$ be the lattice that we defined in §3.1.

For $w\in\mathbf{L}(\varrho_1)$, define $D_3(w,i)$, $1\leqslant i\leqslant m(w)$, to be the connected components of $F^{-n}(\mathbf{B}(w,3\varrho_1))$ that meet $\Lambda(\chi,\varepsilon,k,n;F)$. By Lemma 5.1 and the choice of n_* above, we can check that the restriction of F^n to $D_3(w,i)$ is a diffeomorphism onto $\mathbf{B}(w,3\varrho_1)$, and that $D_3(w,i)$ is contained in $\Lambda(\chi,\varepsilon,k+1,n;F)$. Let $D_1(w,i)$ and $D_2(w,i)$ be the part of $D_3(w,i)$ that F^n maps onto $\mathbf{B}(w,\varrho_1)$ and $\mathbf{B}(w,2\varrho_1)$, respectively. For $\sigma=1,2,3$, let $D_\sigma(w)$ be the union of $D_\sigma(w,i)$ for $1\leqslant i\leqslant m(w)$.

Since the disks $\mathbf{B}(w,\varrho_1)$ for $w\in\mathbf{L}(\varrho_1)$ cover the torus $\mathbf{T}$, we have

$$\mu\circ F^{-n}\leqslant\sum_{w\in\mathbf{L}(\varrho_1)}(\mu\circ F^{-n})|_{\mathbf{B}(w,\varrho_1)}.$$

The function $J_\varrho((\mu\circ F^{-n})|_{\mathbf{B}(w,\varrho_1)})$ is supported on the disk $\mathbf{B}(w,2\varrho_1)$ as $\varrho<\varrho_1$ from the assumption on ϱ. And the intersection multiplicity of the disks $\mathbf{B}(w,2\varrho_1)$ for $w\in\mathbf{L}(\varrho_1)$ is bounded by 10^2 at most. Thus we obtain, by the Schwarz inequality,

$$\begin{aligned}\|\mu\circ F^{-n}\|_\varrho^2 &\leqslant \int\Bigl(\sum_{w\in\mathbf{L}(\varrho_1)} J_\varrho((\mu\circ F^{-n})|_{\mathbf{B}(w,\varrho_1)})(z)\Bigr)^2 d\mathbf{m}(z)\\ &\leqslant 10^2\int\sum_{w\in\mathbf{L}(\varrho_1)} J_\varrho((\mu\circ F^{-n})|_{\mathbf{B}(w,\varrho_1)})(z)^2\, d\mathbf{m}(z)\\ &= 10^2\sum_{w\in\mathbf{L}(\varrho_1)} \|(\mu\circ F^{-n})|_{\mathbf{B}(w,\varrho_1)}\|_\varrho^2.\end{aligned}$$

Since the intersection multiplicity of the regions $D_2(w)$ for $w\in\mathbf{L}(\varrho_1)$ is also bounded by 10^2, we have $\sum_{w\in\mathbf{L}(\varrho_1)}\mu|_{D_2(w)}\leqslant 10^2\mu$ and hence

$$\begin{aligned}\sum_{w\in\mathbf{L}(\varrho_1)}\|\mu|_{D_2(w)}\|_\varrho^2 &= \int\sum_{w\in\mathbf{L}(\varrho_1)} J_\varrho(\mu|_{D_2(w)})(z)^2\, d\mathbf{m}(z)\\ &\leqslant \int(10^2 J_\varrho\mu(z))^2\, d\mathbf{m}(z)\leqslant 10^4\|\mu\|_\varrho^2.\end{aligned}$$

Therefore we can deduce the inequality in the lemma from its localized version:

$$\begin{aligned}\|(\mu\circ F^{-n})|_{\mathbf{B}(w,\varrho_1)}\|_\varrho^2 \leqslant{}& \frac{\mathbf{N}(\chi,\varepsilon,k+1,n;F)\,\|\mu|_{D_2(w)}\|_\varrho^2}{\exp((\chi_c^-+\chi_u^--\chi_c^\Delta-\chi_u^\Delta-\varepsilon)n)}\\ &+\frac{\exp((-2\chi_c^++\varepsilon)n)}{\delta^2}\mu(D_2(w))^2\end{aligned}\tag{59}$$

for $w\in\mathbf{L}(\varrho_1)$, provided that we take the constant n_* so large that $\exp(\varepsilon n_*)>10^6$.

Below we fix $w\in\mathbf{L}(\varrho_1)$ and prove the inequality (59). From the definition of $D_3(w,i)$ and the assumption that μ is supported on $\Lambda(\chi,\varepsilon,k,n;F)$, we have

$$(\mu\circ F^{-n})|_{\mathbf{B}(w,\varrho_1)} = \sum_{i=1}^{m(w)}\mu|_{D_1(w,i)}\circ F^{-n}.$$

Hence the left-hand side of the inequality (59) is written in the form

$$\sum_{1\leqslant i,j\leqslant m(w)}(\mu|_{D_1(w,i)}\circ F^{-n},\mu|_{D_1(w,j)}\circ F^{-n})_\varrho.\tag{60}$$

For $1\leqslant i\leqslant m(w)$, let z_i be the unique point in $D_3(w,i)$ such that $F^n(z_i)=w$, which belongs to $\Lambda(\chi,\varepsilon,k+1,n;F)$. For $1\leqslant i,j\leqslant m(w)$, we write $i\pitchfork j$ if the pair (z_i,z_j) does not belong to the subset $\mathcal{E}(w;\chi,\varepsilon,k+1,n;F)$, that is,

$$\angle(DF^n(\mathbf{E}^u(z_i)),DF^n(\mathbf{E}^u(z_j)))>5H_g\exp((\chi_c^+-\chi_u^-)n+2(k+1)).$$

(See §3.7 for the definition of the set $\mathcal{E}(\,\cdot\,)$.) We split the sum (60) into two parts according to the condition $i\pitchfork j$, and reduce the inequality (59) to the two inequalities

$$\sum_{i\not\pitchfork j}(\mu|_{D_1(w,i)}\circ F^{-n},\mu|_{D_1(w,j)}\circ F^{-n})_\varrho\leqslant\frac{\mathbf{N}(\chi,\varepsilon,k+1,n;F)\,\|\mu|_{D_2(w)}\|_\varrho^2}{\exp((\chi_c^-+\chi_u^--\chi_c^\Delta-\chi_u^\Delta-\varepsilon)n)}$$

and

$$\sum_{i\pitchfork j}(\mu|_{D_1(w,i)}\circ F^{-n},\mu|_{D_1(w,j)}\circ F^{-n})_\varrho\leqslant\frac{\exp((-2\chi_c^++\varepsilon)n)}{\delta^2}\mu(D_2(w))^2.$$

Let $\Sigma_{\not\pitchfork}$ and $\Sigma_{\pitchfork}$ be the sums on the left-hand sides of these two inequalities, respectively.

We prove the first inequality. By the Schwarz inequality, we have

$$\Sigma_{\not\pitchfork}\leqslant\sum_{i\not\pitchfork j}\frac{1}{2}\big(\|\mu|_{D_1(w,i)}\circ F^{-n}\|_\varrho^2+\|\mu|_{D_1(w,j)}\circ F^{-n}\|_\varrho^2\big).$$

Since each term $\|\mu|_{D_1(w,i)}\circ F^{-n}\|_\varrho$ appears at most $2\mathbf{N}(\chi,\varepsilon,k+1,n;F)$ times on the right-hand side, this implies that

$$\Sigma_{\not\pitchfork}\leqslant\mathbf{N}(\chi,\varepsilon,k+1,n;F)\sum_{i=1}^{m(w)}\|\mu|_{D_1(w,i)}\circ F^{-n}\|_\varrho^2.$$

Moreover, we have $\sum_{i=1}^{m(w)}\|\mu|_{D_2(w,i)}\|_\varrho^2\leqslant\|\mu|_{D_2(w)}\|_\varrho^2$. Therefore it is enough to show that

$$\|\mu|_{D_1(w,i)}\circ F^{-n}\|_\varrho^2\leqslant\frac{\|\mu|_{D_2(w,i)}\|_\varrho^2}{\exp(\chi_c^-+\chi_u^--\chi_c^\Delta-\chi_u^\Delta-\varepsilon)}. \tag{61}$$

We show this inequality by using Lemma 6.4. Unfortunately, we cannot apply Lemma 6.4 directly to the measure $\mu|_{D_2(w,i)}$ because some part of its admissible lift may be supported on the part of $\mathbf{AC}((0,\infty))$ that corresponds to very short admissible curves, as a consequence of the restriction. We argue as follows: Observe that F^n brings any C^1-curve with length less than δ in $D_3(w,i)\subset\Lambda(\chi,\varepsilon,k+1,n;F)$ to a curve with length less than ϱ_1 from the assumption on δ and (6), provided that n_* is larger than some constant which depends only on ε, k and the constant C_g in (6). Suppose that an admissible curve γ with length $a\geqslant\delta$ meets $D_2(w,i)$ and that a connected component I of $\gamma^{-1}(D_2(w,i))$ has length less than δ. Then the curve $\gamma|_I$ meets the boundary of $D_2(w,i)$, and hence $F_*^n(\gamma|_I)$ meets the boundary of $\mathbf{B}(w,2\varrho_1)$. From the observation above, $F_*^n(\gamma|_I)$ does not meet $\mathbf{B}(w,\varrho_1)$, and hence $\gamma|_I$ does not meet $D_1(w,i)$. Using this fact, we can construct a measure $\tilde{\mu}$ in $\mathcal{AM}([\delta,\infty))$ that satisfies $\mu|_{D_1(w,i)}\leqslant\tilde{\mu}\leqslant\mu|_{D_2(w,i)}$ by discarding the part of the admissible lift of $\mu|_{D_2(w,i)}$ that is supported on $\mathbf{AC}((0,\delta))$. Note that the observation above also implies that the δ-neighborhood of $D_2(w,i)$ is contained in $D_3(w,i)$, so that $\max\{\#(F^{-n}(z)\cap\mathbf{B}(D_2(w,i),\delta))\,|\,z\in M\}=1$. Now we apply Proposition 6.4 to $\tilde{\mu}$ and $X=D_2(w,i)$. Then the corresponding conclusion and (55) imply (61), provided that n_* is larger than some constant which depends only on ε, k, ϱ_ε, $\varkappa_g$ and I_g.

Next we prove the second inequality. It is enough to show that

$$(\mu|_{D_1(w,i)}\circ F^{-n}, \mu|_{D_1(w,j)}\circ F^{-n})_\varrho \leqslant \delta^{-2}\exp((-2\chi_c^+ +\varepsilon)n)\,\mu(D_2(w,i))\,\mu(D_2(w,j)) \quad (62)$$

for $1\leqslant i,j\leqslant m(w)$ such that $i\pitchfork j$. Both sides of this inequality are linear with respect to $\mu|_{D_2(w,i)}$ and $\mu|_{D_2(w,j)}$. Hence, without loss of generality, we can assume that $\mu|_{D_2(w,i)}$ (resp. $\mu|_{D_2(w,j)}$) has an admissible lift supported on a single element $\{\gamma_i\}\times[0,a_i]$ (resp. $\{\gamma_j\}\times[0,a_j]$) of the partition $\Xi_{\mathbf{AC}}$, and that the curve γ_i (resp. γ_j) is a connected component of the intersection of $D_2(w,i)$ (resp. $D_2(w,j)$) with an admissible curve of length $\geqslant\delta$. From the argument in the proof of the first inequality above, if the length of the curve γ_i (resp. γ_j) is less than δ, it cannot meet $D_1(w,i)$ (resp. $D_1(w,j)$), and hence the inequality (62) is trivial. Thereby, we can also assume that the lengths of γ_i and γ_j, that is, a_i and a_j, are not less than δ.

By the definition of admissible measure and that of the semi-norm $\|\cdot\|_\varrho$, we have

$$\begin{aligned}
&\frac{(\mu|_{D_1(w,i)}\circ F^{-n}, \mu|_{D_1(w,j)}\circ F^{-n})_\varrho}{\mu(D_2(w,i))\,\mu(D_2(w,j))}\\
&\qquad\leqslant \frac{C_g}{a_i a_j(\pi\varrho^2)^2}\int_{\mathbf{T}\times[0,a_i]\times[0,a_j]} \mathbf{1}_\varrho(F^n\circ\gamma_i(t),y)\,\mathbf{1}_\varrho(F^n\circ\gamma_j(s),y)\,d\mathbf{m}(y)\,dt\,ds\\
&\qquad\leqslant C_g\delta^{-2}\varrho^{-2}\int_{[0,a_i]\times[0,a_j]}\mathbf{1}_{2\varrho}(F^n\circ\gamma_i(t),F^n\circ\gamma_j(s))\,dt\,ds.
\end{aligned}$$

We estimate the last term by using the assumption $i\pitchfork j$. From (22), it follows that

$$\angle(DF^n(\mathbf{E}^u(\gamma_i(t))), DF^n(\gamma_i'(t)))\leqslant H_g\exp((\chi_c^+-\chi_u^-)n+2(k+1))$$

for $t\in[0,a_i]$. From Lemma 5.1 (vii), it follows that

$$\angle(DF^n(\mathbf{E}^u(z_i)), DF^n(\mathbf{E}^u(\gamma_i(t))))\leqslant\varkappa_\varepsilon e^{2(k+1)}2\varrho_1\leqslant H_g\exp((\chi_c^+-\chi_u^-)n+2(k+1))$$

for $t\in[0,a_i]$, where the second inequality follows from the definition of ϱ_1, provided that n_* is larger than some constant which depends only on ε, $\varkappa_\varepsilon$ and H_g. Thus we have

$$\angle(DF^n(\mathbf{E}^u(z_i)), DF^n(\gamma_i'(t)))\leqslant 2H_g\exp((\chi_c^+-\chi_u^-)n+2(k+1)) \quad \text{for } t\in[0,a_i],$$

and the same estimate with the index i replaced by j. Therefore the condition $i\pitchfork j$ implies that, for any $t\in[0,a_i]$ and $s\in[0,a_j]$,

$$\angle(DF^n(\gamma_i'(t)), DF^n(\gamma_j'(s)))>H_g\exp((\chi_c^+-\chi_u^-)n+2(k+1)).$$

By simple geometric consideration using this fact, we can see that the part of the curve $F_*^n\gamma_i$ that is within distance 2ϱ from the curve $F_*^n\gamma_j$ has length less than

$$C_g\varrho\exp(-(\chi_c^+-\chi_u^-)n-2(k+1)).$$

Since γ_i and γ_j are admissible curves in $\Lambda(\chi,\varepsilon,k+1,n;F)$, we obtain

$$\begin{aligned}\mathbf{m}_{\mathbf{R}}(\{t\in[0,a_i]\mid d(F^n(\gamma_i(t)),F_*^n\gamma_j)\leqslant 2\varrho\})&\leqslant\frac{C_g\varrho\exp(-(\chi_c^+-\chi_u^-)n-2(k+1))}{\exp(\chi_u^-n-(k+1))}\\&=C_g\varrho\exp(-\chi_c^+n-(k+1))\end{aligned}$$

and the same inequality with the indices i and j exchanged. These facts imply that

$$\int_{[0,a_1]\times[0,a_2]}\mathbf{1}_{2\varrho}(F(\gamma_i(t)),F(\gamma_j(s)))\,dt\,ds\leqslant C_g\varrho^2\exp(-2\chi_c^+n-2(k+1)).$$

Therefore we can conclude (62) by taking the constant n_* larger if necessary. □

6.4. The proof of Theorem 3.21: Part I

We give the proof of Theorem 3.21 in the following three subsections. From this point to the end of this section, we consider the situation assumed in the theorem: Let $\mathbf{X}$ be a finite collection of quadruples $\chi(l)=\{\chi_c^-(l),\chi_c^+(l),\chi_u^-(l),\chi_u^+(l)\}$, $1\leqslant l\leqslant l_0$, satisfying (18), (28), (29) and (30); Let F be a mapping in $\mathcal{U}$ that satisfy the no flat contact condition and the transversality condition on unstable cones for $\mathbf{X}$. The aim of this subsection is to derive the conclusions of Theorem 3.21 from the following proposition:

PROPOSITION 6.6. *Under the assumptions as above, the following claim holds: Let μ_i, $i\geqslant 1$, be a sequence of Borel probability measures on M. We assume that either*

(A) *every μ_i is invariant and has an admissible lift, or*

(B) *$\mu_i=n(i)^{-1}\sum_{j=0}^{n(i)-1}\mathbf{m}_X\circ F^{-j}$ for some subsequence $n(i)\to\infty$, where $\mathbf{m}_X$ is the normalization of the restriction of the Lebesgue measure $\mathbf{m}$ to some Borel subset $X\subset M$ with positive Lebesgue measure.*

Further, we assume that μ_i converges weakly to a Borel probability measure μ_∞ as $i\to\infty$, and that the pair of Lyapunov exponents $(\chi_c(z;F),\chi_u(z;F))$ is contained in the region $|\mathbf{X}|$ for μ_∞-almost every point z. Then, for sufficiently large i, there exists a measure $\nu_i\leqslant\mu_i$ such that

(a) $|\nu_i|>\frac{1}{3}$;

(b) *ν_i is absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$, and the L^2-norm of the density $d\nu_i/d\mathbf{m}$ is bounded by a constant independent of i.*

We assume Proposition 6.6 and prove Theorem 3.21.

Proof of Theorem 3.21. First, note that, if an ergodic invariant measure μ has an admissible lift, and if the pair of Lyapunov exponents $(\chi_c(\mu;F),\chi_u(\mu;F))$ of μ is contained in $|\mathbf{X}|$, then μ is absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$, and hence is a physical measure. This follows immediately from Proposition 6.6 if we set $\mu_i=\mu_\infty=\mu$ in the assumption (A).

We show that there exist at most finitely many ergodic physical measures. Suppose that there exist infinitely many mutually distinct ergodic physical measures μ_i, $i=1,2,\dots$. By taking a subsequence, we can assume that μ_i converges weakly to some measure μ_∞ as $i\to\infty$. We have $\chi_c(\mu_\infty;F)=0$ from Corollary 4.5, Proposition 4.8 and Corollary 3.16. Moreover, we have $\chi_c(z;F)=0$ for μ_∞-almost every point z. In fact, otherwise, there should be an ergodic physical measure $\mu'_\infty\ll\mu_\infty$ with negative central Lyapunov exponent from Lemma 4.6, and hence $\mu_i=\mu'_\infty$ for sufficiently large i from Lemma 4.3, which contradicts the assumption that μ_i are mutually distinct. Since $\lambda_g\leqslant\chi_u(z;F)\leqslant\Lambda_g$ for any point $z\in M$ from the choice of the constants λ_g and Λ_g, the assumption (30) implies that the pair of Lyapunov exponents $(\chi_c(z;F),\chi_u(z;F))$ is contained in $|\mathbf{X}|$ for μ_∞-almost every point z. Therefore we can apply Proposition 6.6 with assumption (A) to the sequence μ_i and conclude that there is a measure $\nu_i\leqslant\mu_i$ for sufficiently large i such that $|\nu_i|>\frac{1}{3}$ and $\|d\nu_i/d\mathbf{m}\|_{L^2(\mathbf{m})}<C$ for a constant C that is independent of i. For these measures ν_i, the Schwarz inequality gives

$$\left(\frac{1}{3}\right)^2<|\nu_i|^2\leqslant\mathbf{m}(\mathcal{B}(\mu_i))\left\|\frac{d\nu_i}{d\mathbf{m}}\right\|^2_{L^2(\mathbf{m})}<C^2\mathbf{m}(\mathcal{B}(\mu_i)).$$

Obviously this contradicts the fact that the basins $\mathcal{B}(\mu_i)$ are mutually disjoint.

Let $\mathcal{B}^0$ be the union of the basins of the ergodic physical measures whose central Lyapunov exponent is neutral. Below we prove that the Lebesgue measure of the subset $X:=M\setminus(\mathcal{B}^-\cup\mathcal{B}^0\cup\mathcal{B}^+)$ is zero. Again the proof is by contradiction. Suppose that the subset X has positive Lebesgue measure. Then, by choosing a subsequence $n(i)\to\infty$ appropriately, we can assume that the sequence of measures $\mu_i=n(i)^{-1}\sum_{j=0}^{n(i)-1}\mathbf{m}_X\circ F^{-j}$ converges to some measure μ_∞ as $i\to\infty$. Note that the measures μ_i are supported on X for $F(X)\subset X$. From Proposition 4.9, we have $\chi_c(z;F)=0$ for μ_∞-almost every point z. Thus the assumption (30) implies that the pair of Lyapunov exponents $(\chi_c(z;F),\chi_u(z;F))$ is contained in $|\mathbf{X}|$ for μ_∞-almost every point z. Each ergodic component of μ_∞ has an admissible lift from Lemma 3.14, and hence it is a physical measure with neutral central Lyapunov exponent from the fact we noted in the beginning. In particular, μ_∞ is supported on $\mathcal{B}^0$. Now apply Proposition 6.6 with assumption (B) to the sequence μ_i, and then let ν_i be those in the corresponding conclusion. Since the density

$\psi_i := d\nu_i/d\mathbf{m}$ has uniformly bounded L^2-norm for sufficiently large i, we can assume that ψ_i converges weakly to some $\psi_\infty \in L^2(\mathbf{m})$, by taking a subsequence of $n(i)$. Note that ψ_∞ is not trivial because

$$(\psi_\infty, 1)_{L^2(\mathbf{m})} = \lim_{i\to\infty} (\psi_i, 1)_{L^2(\mathbf{m})} = \lim_{i\to\infty} |\nu_i| \geqslant \tfrac{1}{3}.$$

On the one hand, we have $\int \psi_i \, d\mu_\infty = 0$ since $\nu_i \leqslant \mu_i$ is supported on $X \subset M \setminus \mathcal{B}^0$. On the other hand, we should have

$$\lim_{i\to\infty} \int \psi_i \, d\mu_\infty \geqslant \lim_{i\to\infty} \int \psi_i \psi_\infty \, d\mathbf{m} = \lim_{i\to\infty} (\psi_i, \psi_\infty)_{L^2} = \|\psi_\infty\|^2_{L^2(\mathbf{m})} > 0$$

because $\psi_\infty \mathbf{m} \leqslant \mu_\infty$. We have arrived at a contradiction.

We have proved that there exists only finitely many ergodic physical measures for F and that the union of basins of them has total Lebesgue measure. The last statement of Theorem 3.21 follows from Proposition 4.7 and the fact that we noted in the beginning of this proof. □

6.5. The proof of Theorem 3.21: Part II

In this subsection, we give the proof of Proposition 6.6, assuming a lemma, Lemma 6.8, whose proof is left to the next subsection. Let μ_i and μ_∞ be those in Proposition 6.6. We put

$$\chi_c^\Delta(l) = \chi_c^+(l) - \chi_c^-(l) \quad \text{and} \quad \chi_u^\Delta(l) = \chi_u^+(l) - \chi_u^-(l) \quad \text{for } 1 \leqslant l \leqslant l_0.$$

To begin with, we fix several constants in the following order:

(K1) Take $0 < \varepsilon < 1$ so small that (43) holds for all the quadruples $\chi \in \mathbf{X}$ and that

$$\lim_{k\to\infty} \liminf_{n\to\infty} \max_{1\leqslant l\leqslant l_0} \frac{\log \mathbf{N}(\chi(l), \varepsilon, k, n; F)}{n(\chi_c^-(l) + \chi_u^-(l) - \chi_c^\Delta(l) - \chi_u^\Delta(l) - 100\varepsilon)} < 1.$$

This is possible from the transversality condition on unstable cones for $\mathbf{X}$.

(K2) Take positive constants ϱ_ε so small and $\varkappa_\varepsilon$ so large that Lemma 5.1 and Lemma 6.4 hold for all the quadruples $\chi \in \mathbf{X}$ and ε above.

(K3) Take a positive constant η so small that

$$10\Lambda_g \eta < \varepsilon \quad \text{and} \quad \eta < 10^{-3}\varepsilon < 10^{-3}.$$

(K4) Take positive constants h_0 and m_0 so large that $h_0 > \Lambda_g > 1$, $m_0 \geqslant n_g$ and

$$\int \min\{0, L(F^n(z); F) + h_0\} \, d\mu(z) > -\frac{\eta}{100} |\mu|$$

for any $\mu \in \mathcal{AM}([1,\infty))$ and $n \geqslant m_0$, where $L(\cdot)$ is the function defined in (17). This is possible from Lemma 3.15 (n_g is the constant we took in §3.4).

(K5) Take a positive constant k_0 such that $k_0 > h_0$ and

$$\mu_\infty\Big(\bigcup_{l=1}^{l_0} \Lambda(\chi(l), \varepsilon, k_0-1, n; F)\Big) > 1 - \frac{\eta}{200 h_0} \quad \text{for any } n > 0.$$

This is possible from Lemma 3.17 and the assumption on μ_∞.

(K6) Take a large positive integer p_0 such that

(a) $\mathbf{N}(\chi(l), \varepsilon, k_0+2, p_0; F) \leqslant \exp((\chi_c^-(l) + \chi_u^-(l) - \chi_c^\Delta(l) - \chi_u^\Delta(l) - 100\varepsilon)p_0)$;

(b) $p_0 > n_*(\chi(l), \varepsilon, k_0+1)$

for $1 \leqslant l \leqslant l_0$, where $n_*(\cdot)$ is given in Lemma 6.5. This is possible from the choice of ε and the fact that $\mathbf{N}(\chi(l), \varepsilon, k, p_0; F)$ is increasing with respect to k.

Hereafter we will never change the constants taken in (K1)–(K5). Note that we can choose the integer p_0 arbitrarily large in the condition (K6) above. In some places below, we shall put additional conditions that p_0 is larger than some numbers that depend only on $\mathbf{X}$, c_g, λ_g, Λ_g, $\varkappa_g$, l_0 and the constants taken in (K1)–(K5).

For a point $z \in M$, we let

$$\mathbf{k}(z) = \min\Big\{k \in \mathbf{Z} \,\Big|\, k \geqslant k_0 \text{ and } z \in \bigcup_{l=1}^{l_0} \Lambda(\chi(l), \varepsilon, k, p_0; F)\Big\} \geqslant k_0$$

and $\mathbf{k}(z) = \infty$ if the set $\{\cdot\}$ above is empty. We also put

$$\mathbf{I}(z) = \begin{cases} 0, & \text{if } \mathbf{k}(z) = k_0, \\ 1, & \text{if } \mathbf{k}(z) > k_0. \end{cases}$$

This is the indicator function of the complement of $\bigcup_{l=1}^{l_0} \Lambda(\chi(l), \varepsilon, k_0, p_0; F)$. Let m be a positive integer and write it in the form $m = q(m)p_0 + d(m)$, where $q(m) = [m/p_0]$, so that $0 \leqslant d(m) < p_0$. We define the subset $\mathcal{R}(m)$ as the set of points $z \in M$ that satisfy

(R1) $\#\{1 \leqslant j \leqslant q \,|\, \mathbf{I}(F^{m-jp_0}(z)) = 1\} < \eta q / 10 h_0$ for $1 \leqslant q \leqslant q(m)$;

(R2) $\sum_{j=1}^{q} (\mathbf{k}(F^{m-jp_0}(z)) - k_0) < \eta q p_0$ for $1 \leqslant q \leqslant q(m)$;

(R3) $\mathbf{k}(z) - k_0 < \eta m$.

The following lemma gives a sufficient condition in order that $\mathcal{R}(m)$, $m = 1, 2, \dots$, are not very small with respect to a measure μ:

LEMMA 6.7. *Let μ be a Borel probability measure μ on M, and n a positive integer such that $n \geqslant 10 p_0$. Assume that*

$$\sum_{j=0}^{n-1} \int |L(F^j(z); F)| \, \mathbf{I}(F^j(z)) \, d\mu(z) < \frac{\eta n}{10} \tag{63}$$

and that

$$\sum_{j=0}^{n-1} \int \mathbf{I}(F^j(z))\, d\mu(z) < \frac{\eta n}{100 h_0}. \tag{64}$$

Then we have $n^{-1}\sum_{m=0}^{n-1} \mu(\mathcal{R}(m)) \geqslant \frac{1}{2}$.

Proof. For $0 \leqslant m < n$, let $\mathcal{Q}_1(m)$, $\mathcal{Q}_2(m)$ and $\mathcal{Q}_3(m)$ be the sets of points z that *violate* the conditions (R1), (R2) and (R3), respectively. We are going to estimate the measures of these subsets by using Lemma 4.14. First we give the estimate on the subset $\mathcal{Q}_1(m)$ for $0 \leqslant m < n$. If $z \in \mathcal{Q}_1(m)$, we have

$$\sum_{j=1}^{q} \mathbf{I}(F^{m-jp_0}(z)) \geqslant \frac{\eta q}{10 h_0}$$

for some $1 \leqslant q < q(m)$. Using Lemma 4.14 with the assumption (64), we obtain

$$\begin{aligned}\sum_{m=0}^{n-1} \mu(\mathcal{Q}_1(m)) &= \sum_{d=1}^{p_0} \sum_{j=0}^{[(n-d)/p_0]} \mu(\mathcal{Q}_1((n-d)-jp_0)) \\ &\leqslant \sum_{d=1}^{p_0} \left(\frac{10 h_0}{\eta} \sum_{j=0}^{[(n-d)/p_0]} \int \mathbf{I}(F^{(n-d)-jp_0}(z))\, d\mu(z) \right) \leqslant \frac{n}{10}.\end{aligned}$$

Next we give the estimate on the union $\mathcal{Q}_2(m) \cup \mathcal{Q}_3(m)$. Let us put

$$\psi(z) = (|L(z;F)| + 5\Lambda_g)\, \mathbf{I}(z).$$

We claim that

$$\mathbf{k}(z) - k_0 \leqslant \sum_{j=0}^{p_0-1} \psi(F^j(z)) \quad \text{for } z \in M. \tag{65}$$

For a point z, take the smallest integer $0 \leqslant p < p_0$ such that $\mathbf{k}(F^p(z)) = k_0$, and set $p = p_0$ if there are no such integers. If $p = 0$, the inequality (65) is trivial. So we assume $p > 0$. In the case $0 < p < p_0$, we choose an integer $1 \leqslant l \leqslant l_0$ so that $\Lambda(\chi(l), \varepsilon, k_0, p_0; F)$ contains $F^p(z)$. In the case $p = p_0$, we choose $1 \leqslant l \leqslant l_0$ arbitrarily. For $0 \leqslant i < i' \leqslant p$ and $v \in \mathbf{S}^u(F^i(z))$, we have the obvious estimates

$$\sum_{j=i}^{i'-1} L(F^j(z);F) \leqslant \log |D^* F^{i'-i}(v)| \leqslant \Lambda_g (i'-i),$$

$$-\Lambda_g \leqslant -c_g \leqslant \log |D_* F^{i'-i}(v)| \leqslant \Lambda_g (i'-i).$$

Using these estimates and the fact that $F^p(z)\in\Lambda(\chi(l),\varepsilon,k_0,p_0;F)$ in the case $p<p_0$, we can check that z belongs to $\Lambda(\chi(l),\varepsilon,k,p_0;F)$ for

$$k=k_0+\left[\sum_{j=0}^{p-1}(|L(F^j(z);F)|+3\Lambda_g+\varepsilon)\right]+1.$$

This implies (65).

If a point z belongs to $\mathcal{Q}_2(m)$ or $\mathcal{Q}_3(m)$ for $p_0\leqslant m<n$, we have, from (65),

$$\sum_{j=m'}^{m-1}\psi(F^j(z))\geqslant\eta(m-m')\quad\text{for some } 0\leqslant m'<m.$$

As we took h_0 so that $h_0>\Lambda_g$, the assumptions (63) and (64) imply

$$\sum_{j=0}^{n-1}\int\psi(F^j(z))\,d\mu(z)\leqslant\frac{\eta n}{5}.$$

Therefore, by using Lemma 4.14, we can obtain

$$\sum_{m=p_0}^{n-1}\mu(\mathcal{Q}_2(m)\cup\mathcal{Q}_3(m))\leqslant\frac{n}{5}.$$

Note that we have $\sum_{m=0}^{p_0}\mu(\mathcal{Q}_2(m)\cup\mathcal{Q}_3(m))\leqslant p_0\leqslant\frac{1}{10}n$, as we assume that $n\geqslant 10p_0$ in the lemma. Since $\mathcal{R}(m)$ is the complement of $\mathcal{Q}_1(m)\cup\mathcal{Q}_2(m)\cup\mathcal{Q}_3(m)$, we can obtain the lemma from the estimates above. □

The following lemma is the key step in the proof of Proposition 6.6:

LEMMA 6.8. *Let μ be a Borel finite measure on M, and n a non-negative integer. If μ has an admissible lift $\tilde{\mu}$ such that $\tilde{\mu}\circ F_*^{-i}$ belongs to* $\mathbf{AM}([\exp(-\eta n),\infty))$ *for $0\leqslant i<n$, then we have*

$$\|\mu|_{\mathcal{R}(n)}\circ F^{-n}\|_{\varrho}<C|\mu|+C\exp(-\varepsilon n)\,\|\mu\|_{\varrho\exp(-10\eta n)}$$

for $0<\varrho\leqslant\exp(-10\Lambda_g p_0)$, where $C>0$ is a constant that does not depend on the measure μ nor the integer n.

Remark. Actually, the constant $C>0$ above depends only on ε, p_0, c_g and Λ_g.

We give the proof of this lemma in the next subsection. Below we assume this lemma and complete the proof of Proposition 6.6.

Proof of Proposition 6.6. First consider the case where the assumption (A) holds. From the choice of k_0, we have

$$\mu_i\Big(\bigcup_{l=1}^{l_0}\Lambda(\chi(l),\varepsilon,k_0,p_0;F)\Big)>1-\frac{\eta}{100h_0},$$

or, in other words,

$$\int \mathbf{I}(z)\,d\mu_i < \frac{\eta}{100h_0}$$

for sufficiently large i, because $\Lambda(\chi(l),\varepsilon,k_0,p_0;F)$ contains an open neighborhood of the compact subset $\Lambda(\chi(l),\varepsilon,k_0-1,p_0;F)$. The measure μ_i belongs to $\mathcal{AM}([1,\infty))$ from Corollary 3.8. Thus, it follows from the choice of h_0 that

$$\int \min\{0, L(z;F)+h_0\}\,d\mu_i(z) > -\frac{\eta}{100}.$$

Hence

$$\int |L(z;F)|\,\mathbf{I}(z)\,d\mu_i(z) < h_0\frac{\eta}{100h_0}+\frac{\eta}{100} < \frac{\eta}{10}.$$

Now we can apply Lemma 6.7 to the invariant measure μ_i for sufficiently large i, and obtain

$$\frac{1}{n}\sum_{j=0}^{n-1}\mu_i(\mathcal{R}(j)) \geqslant \frac{1}{2} \quad \text{for } n\geqslant 10p_0.$$

We put

$$\nu_{i,n} = \frac{1}{n}\sum_{j=0}^{n-1}\mu_i|_{\mathcal{R}(j)}\circ F^{-j} \leqslant \mu_i \quad \text{for } n\geqslant 1,$$

so that $|\nu_{i,n}|\geqslant\frac{1}{2}$ for $n\geqslant 10p_0$. Obviously the measure μ_i has an admissible lift that satisfies the assumption of Lemma 6.8 for any $n\geqslant 0$. Thus it holds that

$$\|\nu_{i,n}\|_\varrho \leqslant \frac{1}{n}\sum_{j=0}^{n-1}\|\mu_i|_{\mathcal{R}(j)}\circ F^{-j}\|_\varrho \leqslant C+\frac{C}{n}\sum_{j=0}^{n-1}\exp(-\varepsilon j)\,\|\mu_i\|_{\varrho\exp(-10\eta j)}$$

for $0<\varrho\leqslant\exp(-10\Lambda_g p_0)$. This, together with (54) and the choice of η, implies that $\limsup_{n\to\infty}\|\nu_{i,n}\|_\varrho\leqslant C$. Let ν_i be a weak limit point of the sequence $\nu_{i,n}$, $n=1,2,\ldots$. Then it holds that $\nu_i\leqslant\mu_i$ and $|\nu_i|\geqslant\frac{1}{2}$. Also we have $\|\nu_i\|_\varrho\leqslant C$ for $0<\varrho\leqslant\exp(-10\Lambda_g p_0)$ from Lemma 6.3. From Lemma 6.2, this implies that ν_i is absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$, and the density satisfies $\|d\nu_i/d\mathbf{m}\|_{L^2(\mathbf{m})}\leqslant C$. Thus the measures ν_i satisfy the conditions in Proposition 6.6.

Next we consider the case where the assumption (B) holds. Let $n_0=n_0(F)>n_g$ be as in the definition of the no flat contact condition. Let X and $\mathbf{m}_X$ be as in the assumption (B). Using Lemma 3.12, we can find a small positive number $b>0$ and a probability measure $\omega'\in\mathcal{AM}([b,\infty))$ such that

(1) $|\mathbf{m}_X-\omega'|<10^{-3}\eta/h_0$;

(2) $\omega'\circ F^{-n_0}$ is absolutely continuous with respect to the Lebesgue measure $\mathbf{m}$;

(3) the density of the measure $\omega'\circ F^{-n_0}$, $d(\omega'\circ F^{-n_0})/d\mathbf{m}$, is square integrable.

Remark. In the third condition above, we do not care how large the L^2-norm is.

We put $\omega = \omega' \circ F^{-n_0}$ and

$$\mu_i' = \frac{1}{n(i)} \sum_{j=0}^{n(i)-1} \omega \circ F^{-j} \quad \text{for } i = 1, 2, \dots.$$

Then, for sufficiently large i, we have $|\mu_i - \mu_i'| < 10^{-3}\eta / h_0$ and hence

$$\mu_i' \Big(\bigcup_{l=1}^{l_0} \Lambda(\chi(l), \varepsilon, k_0, p_0; F) \Big) > 1 - \frac{\eta}{100 h_0}, \quad \text{that is,} \quad \int \mathbf{I}(z)\, d\mu_i' < \frac{\eta}{100 h_0}$$

from the choice of k_0. From Corollary 3.8, $\omega \circ F^{-j}$ belongs to $\mathcal{AM}([1, \infty))$ for sufficiently large j. Thus we have

$$\int |L(z; F)|\, \mathbf{I}(z)\, d\mu_i'(z) < h_0 \frac{\eta}{100 h_0} + \frac{\eta}{100} < \frac{\eta}{10}$$

for sufficiently large i, from the choice of h_0. Now we can apply Lemma 6.7 to $\mu = \omega$ and $n = n(i)$ in order to obtain

$$\frac{1}{n(i)} \sum_{m=0}^{n-1} \omega(\mathcal{R}(m)) \geqslant \frac{1}{2}$$

for sufficiently large i. Let $\widetilde{\omega}'$ be an admissible lift of ω' that belongs to $\mathbf{AM}([b, \infty))$, and put $\widetilde{\omega} = \widetilde{\omega}' \circ F_*^{-n_0}$. Then $\widetilde{\omega}$ is an admissible lift of ω. Take a large positive integer n_1 that satisfies $\exp(-\eta n_1) < b \exp(-c_g)$. From Lemma 3.7, the measures $\widetilde{\omega} \circ F_*^{-i} = \widetilde{\omega}' \circ F_*^{-i-n_0}$ for $i \geqslant 0$ belongs to $\mathbf{AM}([\exp(-\eta n), \infty))$, provided that $n \geqslant n_1$. Thus we can apply Lemma 6.8 to ω, and obtain

$$\|\omega|_{\mathcal{R}(n)} \circ F^{-n}\|_\varrho < C|\omega| + C \exp(-\varepsilon n) \|\omega\|_{\varrho \exp(-10\eta n)}$$

for $0 < \varrho \leqslant \exp(-10\Lambda_g p_0)$ and $n \geqslant n_1$. We put

$$\nu_i' = \frac{1}{n(i)} \sum_{j=n_1}^{n(i)-1} \omega|_{\mathcal{R}(j)} \circ F^{-j} \leqslant \mu_i', \quad i = 1, 2, \dots .$$

Then, for sufficiently large i, we have $|\nu_i'| \geqslant \frac{2}{5}$ and

$$\|\nu_i'\|_\varrho \leqslant C + \frac{C}{n(i)} \sum_{j=n_1}^{n(i)-1} \exp(-\varepsilon j) \|\omega\|_{\varrho \exp(-10\eta j)}$$

for $0 < \varrho \leqslant \exp(-10\Lambda_g p_0)$. Letting $\varrho \to +0$ in the last inequality, we obtain

$$\left\| \frac{d\nu_i'}{d\mathbf{m}} \right\|_{L^2} \leqslant C + \left(\frac{C}{n(i)} \sum_{j=n_1}^{n(i)-1} \exp(-\varepsilon j) \right) \left\| \frac{d\omega}{d\mathbf{m}} \right\|_{L^2}$$

by Lemma 6.2. Since we have $|\mu_i' - \mu_i| < 10^{-2}$ and $\nu_i' \leqslant \mu_i'$, we can find a Borel measure ν_i such that $\nu_i \leqslant \nu_i'$, $\nu_i \leqslant \mu_i$, $|\nu_i| > \frac{1}{3}$ and $\|d\nu_i / d\mathbf{m}\|_{L^2} \leqslant 2C$ for sufficiently large i. The measures ν_i satisfy the conditions in Proposition 6.6. $\square$

6.6. The proof of Theorem 3.21: Part III

In this subsection, we give the proof of Lemma 6.8 and complete the proof of Theorem 3.21. Let n, μ and $\tilde{\mu}$ be as in Lemma 6.8. Recall the mapping $\Pi: \mathbf{A}((0,\infty))\to M$ and the commutative relation (14) in §3.4. Below we divide the measure $\tilde{\mu}$ into many parts, so that we can evaluate the semi-norms of their images under the mapping $\Pi\circ F_*^n$ by the two inequalities we gave in §6.3.

We write the integer n in the form $n=q(n)p_0+d(n)$, where $q(n)=[n/p_0]$, so that $0\leqslant d(n)<p_0$. For integers $-1\leqslant q\leqslant q(n)$, we put

$$\tau(q)=\begin{cases} qp_0+d(n) & \text{for } 0\leqslant q\leqslant q(n),\\ 0 & \text{for } q=-1,\end{cases}$$

so that $\tau(q(n))=n$, and we also put

$$\delta(q)=\begin{cases} \exp(-4\eta(n-\tau(q))-7\Lambda_g p_0-c_g) & \text{for } 0\leqslant q\leqslant q(n),\\ \exp(-4\eta n-7\Lambda_g p_0) & \text{for } q=-1.\end{cases}$$

Fix a number $0<\varrho\leqslant\exp(-10\Lambda_g p_0)$ arbitrarily and put

$$\varrho(q)=\varrho\exp(-10\eta(n-\tau(q)))\quad \text{for } -1\leqslant q\leqslant q(n).$$

We put $W=\mathbf{AC}([\exp(-\eta n),\infty))$, so that $\tilde{\mu}\circ F_*^{-i}$ for $0\leqslant i\leqslant n$ are supported on W, by assumption.

We begin with constructing measurable partitions $\xi(q)$, $-1\leqslant q\leqslant q(n)$, of the space W such that:

(Ξ1) $\xi(q)$ subdivides the partition $\Xi_{\mathbf{AC}}$ on W, which is defined in §3.5. And $\xi(q)$ is increasing with respect to q, that is, $\xi(q+1)$ subdivides $\xi(q)$.

(Ξ2) Each element of the partition $\xi(q)$ is of the form $\{\gamma\}\times J$, where γ is an admissible curve in $\mathcal{AC}(a)$ with $a\geqslant\exp(-\eta n)$ and J is an interval in $[0,a]$ such that $\delta(q)\leqslant|F_*^{\tau(q)}(\gamma|_J)|\leqslant 2\delta(q)$.

The construction is easily done by induction on q. Since $\delta(-1)<\exp(-\eta n)$, we can construct a partition $\xi(-1)$ that satisfies (Ξ1) and (Ξ2) by subdividing the partition $\Xi_{\mathbf{AC}}$ on W. Let $0\leqslant q\leqslant q(n)$ and suppose that we have constructed the partitions $\xi(j)$ for $-1\leqslant j<q$. For each element $\{\gamma\}\times J$ of $\xi(q-1)$, the length of the curve $F_*^{\tau(q)}(\gamma|_J)$ is not less than

$$\delta(q-1)\exp(\lambda_g(\tau(q)-\tau(q-1))-c_g)>\delta(q),$$

provided that we take the constant p_0 so large that $(\lambda_g-4\eta)p_0>c_g$. (Recall the remark on the choice of the constant p_0 in the last subsection.) Hence we can construct the partition $\xi(q)$ satisfying (Ξ1) and (Ξ2) by subdividing $\xi(q-1)$.

A Borel measurable subset in W is said to be a $\xi(q)$-subset if it is a union of elements of $\xi(q)$. Note that, if Y is a $\xi(q)$-subset, the measure $\tilde{\mu}|_Y \circ F_*^{-\tau(q)} \circ \Pi^{-1}$ is contained in $\mathcal{AM}([\delta(q), 2\delta(q)])$ by the condition ($\Xi 2$).

For $-1 \leqslant q \leqslant q(n)$ and an element $P = \{\gamma\} \times J$ of the partition $\xi(q)$, we define

$$\mathbf{k}_q(P) := \min\{\mathbf{k}(F^{\tau(q)}(\gamma(t))) \mid t \in J\} \geqslant k_0,$$

where $\mathbf{k}(\cdot)$ was defined in the last subsection. For simplicity, we put

$$\|\tilde{\nu}\|_\varrho := \|\tilde{\nu} \circ \Pi^{-1}\|_\varrho \quad \text{for a measure } \tilde{\nu} \text{ on } W.$$

The following result is a consequence of the two inequalities in §6.3:

SUBLEMMA 6.9. *Let Y be a $\xi(q)$-subset in W for some $-1 \leqslant q \leqslant q(n)$, and let k be an integer such that*

$$k_0 \leqslant k \leqslant k_0 + \eta(n - \tau(q)). \tag{66}$$

If $\mathbf{k}_q(P) \leqslant k$ for all elements $P \in \xi(q)$ that are contained in Y, we have

$$\|\tilde{\mu}|_Y \circ F_*^{-\tau(q+1)}\|_{\varrho(q+1)} \leqslant \exp(10\Lambda_g p_0 + 6(k - k_0)) \|\tilde{\mu}|_Y \circ F_*^{-\tau(q)}\|_{\varrho(q)}.$$

Moreover, if $k = k_0$ and $q \geqslant 0$ in addition, we have either

$$\|\tilde{\mu}|_Y \circ F_*^{-\tau(q+1)}\|_{\varrho(q+1)} \leqslant \exp(-48\varepsilon p_0) \|\tilde{\mu}|_Y \circ F_*^{-\tau(q)}\|_{\varrho(q)}$$

or

$$\|\tilde{\mu}|_Y \circ F_*^{-\tau(q+1)}\|_{\varrho(q+1)} \leqslant \delta(q)^{-1} \exp(3\Lambda_g p_0) \tilde{\mu}(Y).$$

Proof. We put $p = \tau(q+1) - \tau(q) \leqslant p_0$. So p is smaller than p_0 only if $q = -1$. By assumption, we can divide the subset Y into $\xi(q)$-subsets $Y(l)$, $1 \leqslant l \leqslant l_0$, such that $\Pi \circ F_*^{\tau(q)}(P) \cap \Lambda(\chi(l), \varepsilon, k, p_0; F) \neq \varnothing$ for each $P \in \xi(q)$ that is contained in $Y(l)$. The measures $\tilde{\mu}|_{Y(l)} \circ F^{-\tau(q)} \circ \Pi^{-1}$ belong to $\mathcal{AM}([\delta(q), \infty))$, as we noted above.

We prove the first claim. By using (66) and (30), we can check that

$$2\delta(q) \leqslant \varkappa_g^{-1} \varrho_\varepsilon \exp((\chi_c^-(l) - \chi_u^+(l) - 5\varepsilon) p_0 - 4k),$$

provided that p_0 is larger than some constant that depends only on k_0, ϱ_ε, $\varkappa_g$ and Λ_g. This and the claims (v) and (vi) of Lemma 5.1 imply that the subset $\Pi \circ F_*^{\tau(q)}(Y(l))$ is contained in $\Lambda(\chi(l), \varepsilon, k+1, p_0; F)$, and hence is contained in $\Lambda(\chi(l), \varepsilon, k+1+\varepsilon p_0, p; F)$ even in the case $p < p_0$, by (21).

For simplicity, we put

$$\delta := 10\varkappa_g \varrho(q+1)\exp(-\chi_c^-(l)p+k+\varepsilon p_0+2).$$

We can check that

$$\delta < \varkappa_g^{-1}\varrho_\varepsilon \exp((\chi_c^- - \chi_u^+ - 5\varepsilon)p - 4(k+1+\varepsilon p_0)), \tag{67}$$

$$\varrho(q) < \delta < \delta(q), \tag{68}$$

$$0 < \varrho(q+1) \leqslant \frac{\varrho_\varepsilon \exp((\chi_c^-(l)-5\varepsilon)p-3(k+2+\varepsilon p_0))}{10\varkappa_g^2}, \tag{69}$$

by (66) and (30), provided that p_0 is larger than some constant which depends only on k_0, ϱ_ε, $\varkappa_g$, c_g and Λ_g. The subset $\Pi\circ F_*^{\tau(q)}(Y(l))$ is contained in $\Lambda(\chi(l),\varepsilon,k+1+\varepsilon p_0,p;F)$ as we noted, so the claims (v) and (vi) of Lemma 5.1 and the inequality (67) imply that the δ-neighborhood of $\Pi\circ F_*^{\tau(q)}(Y(l))$ is contained in $\Lambda(\chi(l),\varepsilon,k+2+\varepsilon p_0,p;F)$. From Corollary 5.2, it follows that

$$\max_{w\in M} \#(F^{-p}(w)\cap \mathbf{B}(\Pi\circ F_*^{\tau(q)}(Y(l)),\delta)) < \exp(6\Lambda_g p_0+6k),$$

provided that p_0 is larger than some constant that depends only on $\varkappa_\varepsilon$ and Λ_g. Now we can apply Lemma 6.4 and obtain

$$\begin{aligned}\|\tilde\mu|_{Y(l)}\circ F_*^{-\tau(q+1)}\|_{\varrho(q+1)}^2 &\leqslant I_g \exp(16\Lambda_g p_0+6(k+\varepsilon p_0+1)+6k)\,\|\tilde\mu|_{Y(l)}\circ F_*^{-\tau(q)}\|_\delta^2\\ &\leqslant l_0^{-2}\exp(20\Lambda_g p_0+12(k-k_0))\,\|\tilde\mu|_Y\circ F_*^{-\tau(q)}\|_{\varrho(q)}^2\end{aligned}$$

using (55), provided that p_0 is larger than some constant which depends only on I_g, k_0, l_0 and Λ_g. Summing up the square root of both sides over $1\leqslant l\leqslant l_0$, we obtain the first claim.

We prove the second claim by using Lemma 6.5. Note that $\Pi\circ F_*^{\tau(q)}(Y(l))$ is contained in $\Lambda(\chi(l),\varepsilon,k_0+1,p_0;F)$ in this case, by the argument above. We can check that

$$\varrho(q+1)\exp((-\chi_c^-(l)+\varepsilon)p_0) < \delta(q) < \exp((\chi_c^-(l)-2\chi_u^+(l)-3\varepsilon)p_0),$$

provided that p_0 is larger than some constant which depends only on c_g and Λ_g. Recall that we took p_0 so large that $p_0\geqslant n_*(\chi(l),\varepsilon,k_0+1)$ in the condition (K6). Hence we can apply Lemma 6.5 and obtain

$$\begin{aligned}\|\tilde\mu|_{Y(l)}\circ F_*^{-\tau(q+1)}\|_{\varrho(q+1)}^2 &\leqslant \exp(-98\varepsilon p_0)\,\|\tilde\mu|_{Y(l)}\circ F_*^{-\tau(q)}\|_{\varrho(q+1)}^2\\ &\quad+\delta(q)^{-2}\exp((-2\chi_c^+(l)+2\varepsilon)p_0)\,\tilde\mu(Y(l))^2,\end{aligned}$$

where we used the condition (K6) (a) in the choice of p_0. This implies that

$$\|\tilde{\mu}|_{Y(l)} \circ F_*^{-\tau(q+1)}\|_{\varrho(q+1)} \leqslant \exp(-49\varepsilon p_0)\|\tilde{\mu}|_Y \circ F_*^{-\tau(q)}\|_{\varrho(q+1)} + \delta(q)^{-1}\exp((-\chi_c^+(l)+\varepsilon)p_0)\tilde{\mu}(Y).$$

Summing up both sides for $1 \leqslant l \leqslant l_0$ and using (55), we conclude that

$$\|\tilde{\mu}|_Y \circ F_*^{-\tau(q+1)}\|_{\varrho(q+1)} \leqslant C_0 l_0 \exp(-49\varepsilon p_0)\|\tilde{\mu}|_Y \circ F_*^{-\tau(q)}\|_{\varrho(q)} + l_0\delta(q)^{-1}\exp((2\Lambda_g+\varepsilon)p_0)\tilde{\mu}(Y).$$

The second claim follows from this inequality, provided that p_0 is larger than some constant that depends only on l_0, Λ_g and ε. □

For integers $-1 \leqslant q' \leqslant q \leqslant q(n)$, let $\mathcal{K}(q', q)$ be the set of sequences $\sigma = \{\sigma_j\}_{j=q'}^{q-1}$ of $q-q'$ integers that satisfy

$$0 \leqslant \sigma_j \leqslant \eta(n-\tau(j)) \quad \text{for } q' \leqslant j < q. \tag{70}$$

In the case $q'=q$, we say that $\mathcal{K}(q', q) = \mathcal{K}(q, q)$ consists of one empty sequence, which is denoted by $\varnothing_q$. We put

$$\mathcal{K}(q) = \bigcup \{\mathcal{K}(q', q) \mid -1 \leqslant q' \leqslant q\}$$

for $0 \leqslant q \leqslant q(n)$. Below we construct subsets $\mathcal{D}(\sigma)$ in W for $\sigma \in \bigcup_{q=-1}^{q(n)} \mathcal{K}(q)$ so that the following conditions hold:

(D1) $\mathcal{D}(\sigma)$ for $\sigma \in \mathcal{K}(q)$ are mutually disjoint $\xi(q-1)$-subsets.

(D2) The union of $\mathcal{D}(\sigma)$ for all $\sigma \in \mathcal{K}(q)$ contains the subset $\Pi^{-1}(\mathcal{R}(n)) \cap W$.

(D3) For $-1 \leqslant q' < q \leqslant q(n)$ and $\sigma = \{\sigma_j\}_{j=q'}^{q-1} \in \mathcal{K}(q', q)$, we have

$$\|\tilde{\mu}|_{\mathcal{D}(\sigma)} \circ F_*^{-\tau(q)}\|_{\varrho(q)} \leqslant \exp(10\Lambda_g p_0 + 6\sigma_{q-1})\|\tilde{\mu}|_{\mathcal{D}(\sigma')} \circ F_*^{-\tau(q-1)}\|_{\varrho(q-1)},$$

where $\sigma' = \{\sigma_j\}_{j=q'}^{q-2} \in \mathcal{K}(q', q-1)$ (so $\sigma' = \varnothing_{q'}$ if $q' = q-1$). Further, it holds that

$$\|\tilde{\mu}|_{\mathcal{D}(\sigma)} \circ F_*^{-\tau(q)}\|_{\varrho(q)} \leqslant \exp(-48\varepsilon p_0)\|\tilde{\mu}|_{\mathcal{D}(\sigma')} \circ F_*^{-\tau(q-1)}\|_{\varrho(q-1)}$$

in the case where $q \geqslant 1$ and $\sigma_{q-1} = 0$.

(D4) For the empty sequence $\varnothing_q \in \mathcal{K}(q, q)$ for $q \geqslant 0$, we have

$$\|\tilde{\mu}|_{\mathcal{D}(\varnothing_q)} \circ F_*^{-\tau(q)}\|_{\varrho(q)} \leqslant \delta(q-1)^{-1}\exp(3\Lambda_g p_0)\tilde{\mu}(\mathcal{D}(\varnothing_q)).$$

The construction is done by induction on q. For the case $q=-1$, we just put $\mathcal{D}(\varnothing_{-1})=W$. For the case $q=0$, we have to define $\mathcal{D}(\sigma)$ for $\sigma=\varnothing_0\in\mathcal{K}(0,0)$ and $\sigma=\{\sigma_{-1}\}\in\mathcal{K}(-1,0)$, where $0\leqslant\sigma_{-1}\leqslant\eta n$ from (70). We let $\mathcal{D}(\varnothing_0)$ be the empty set and put

$$\mathcal{D}(\{\sigma_{-1}\})=\bigcup\{P\in\xi(-1)\mid \mathbf{k}_{(-1)}(P)=k_0+\sigma_{-1}\}\quad\text{for } 0\leqslant\sigma_{-1}\leqslant\eta n.$$

Then the conditions (D1) and (D4) obviously hold. The condition (D2) follows from the condition (R3) in the definition of the subset $\mathcal{R}(n)$. The first claim of Sublemma 6.9 implies that the condition (D3) also holds.

Next, let $q\geqslant1$ and suppose that we have defined $D(\sigma)$ for $\sigma\in\mathcal{K}(q-1)$ so that the conditions (D1)–(D4) hold for them. Consider an element $\sigma=\{\sigma_j\}_{j=q'}^{q-1}$ in $\mathcal{K}(q',q)$ with $q'<q$ and put $\sigma'=\{\sigma_j\}_{j=q'}^{q-2}\in\mathcal{K}(q',q-1)$. Let us set

$$\mathcal{D}_*(\sigma)=\bigcup\{P\in\xi(q-1)\mid P\subset\mathcal{D}(\sigma')\text{ and }\mathbf{k}_{q-1}(P)=k_0+\sigma_{q-1}\}. \tag{71}$$

In the case $\sigma_{q-1}>0$, we put $\mathcal{D}(\sigma)=\mathcal{D}_*(\sigma)$. In the case $\sigma_{q-1}=0$, we define $\mathcal{D}(\sigma)$ in the following manner: From the second claim of Sublemma 6.9, we have either

$$\|\tilde{\mu}|_{\mathcal{D}_*(\sigma)}\circ F_*^{-\tau(q)}\|_{\varrho(q)}\leqslant\exp(-48\varepsilon p_0)\|\tilde{\mu}|_{\mathcal{D}_*(\sigma)}\circ F_*^{-\tau(q-1)}\|_{\varrho(q-1)} \tag{72}$$

or

$$\|\tilde{\mu}|_{\mathcal{D}_*(\sigma)}\circ F_*^{-\tau(q)}\|_{\varrho(q)}\leqslant\delta(q-1)^{-1}\exp(3\Lambda_g p_0)\,\tilde{\mu}(\mathcal{D}_*(\sigma)). \tag{73}$$

We let

$$\mathcal{D}(\sigma)=\begin{cases}\mathcal{D}_*(\sigma), & \text{when (72) holds,}\\ \varnothing, & \text{otherwise.}\end{cases}$$

Finally we define $\mathcal{D}(\varnothing_q)$ as the union of $\mathcal{D}_*(\sigma)$ for the sequences $\sigma=\{\sigma_j\}_{j=q'}^{q-1}$ in $\bigcup_{-1<q'<q}\mathcal{K}(q',q)=\mathcal{K}(q)\setminus\{\varnothing_q\}$ such that $\sigma_{q-1}=0$ and such that (72) does *not* hold. As a consequence of this definition, the condition (D4) holds for the empty sequence $\varnothing_q$. The condition (D1) obviously holds. We can check the condition (D2) by using the condition (R2) in the definition of the subset $\mathcal{R}(n)$. The condition (D3) follows from the first claim of Sublemma 6.9 and the construction above. We have finished the definition of the subsets $\mathcal{D}(\sigma)$.

For $-1\leqslant q'\leqslant q(n)$, let $\mathcal{K}_*(q')$ be the set of sequences $\sigma=\{\sigma_j\}_{j=q'}^{q(n)-1}$ in $\mathcal{K}(q',q(n))$ that satisfy the conditions

$$|\sigma|_0\leqslant\eta(q(n)-q')\quad\text{and}\quad|\sigma|_1\leqslant2\eta(q(n)-q')p_0,$$

where

$$|\sigma|_0:=\#\{q'\leqslant j<q(n)\mid j\geqslant0\text{ and }\sigma_j>0\}\quad\text{and}\quad|\sigma|_1:=\sum_{j=q'}^{q(n)-1}\sigma_j.$$

Then, from the definition of the subsets $\mathcal{R}(n)$ and $\mathcal{D}(\sigma)$, we have

$$\Pi^{-1}(\mathcal{R}(n))\cap W\subset \bigcup_{q'=-1}^{q(n)} \bigcup_{\sigma\in\mathcal{K}_*(q')} \mathcal{D}(\sigma)$$

and hence

$$\tilde{\mu}|_{\Pi^{-1}(\mathcal{R}(n))}\circ F_*^{-n} \leqslant \sum_{q'=-1}^{q(n)} \sum_{\sigma\in\mathcal{K}_*(q')} \tilde{\mu}|_{\mathcal{D}(\sigma)}\circ F_*^{-n}.$$

For each $\sigma=\{\sigma_j\}_{j=q'}^{q(n)-1}$ in $\mathcal{K}_*(q')$ with $q'\geqslant 0$, we can obtain

$$\begin{aligned}\|\tilde{\mu}|_{\mathcal{D}(\sigma)}\circ F_*^{-n}\|_{\varrho} &= \|\tilde{\mu}|_{\mathcal{D}(\sigma)}\circ F_*^{-\tau(q(n))}\|_{\varrho(q(n))}\\ &\leqslant \exp(10\Lambda_g p_0|\sigma|_0+6|\sigma|_1-48\varepsilon(q(n)-q'-|\sigma|_0)p_0)\,\|\tilde{\mu}|_{\mathcal{D}(\varnothing_{q'})}\circ F_*^{-\tau(q')}\|_{\varrho(q')}\end{aligned}$$

from the condition (D3), and hence

$$\|\tilde{\mu}|_{\mathcal{D}(\sigma)}\circ F_*^{-n}\|_{\varrho} \leqslant \exp(-45\varepsilon(q(n)-q')p_0+11\Lambda_g p_0+c_g)\,|\tilde{\mu}|$$

from the condition (D4) and the choice of η. Similarly, for $\sigma=\{\sigma_j\}_{j=-1}^{q(n)-1}$ in $\mathcal{K}_*(-1)$, we can obtain

$$\|\tilde{\mu}|_{\mathcal{D}(\sigma)}\circ F_*^{-n}\|_{\varrho} \leqslant \exp(10\Lambda_g p_0(|\sigma|_0+1)+6|\sigma|_1-48\varepsilon(q(n)-|\sigma|_0)p_0)\,\|\tilde{\mu}\|_{\varrho(-1)}$$

and hence

$$\|\tilde{\mu}|_{\mathcal{D}(\sigma)}\circ F_*^{-n}\|_{\varrho} \leqslant \exp(-45\varepsilon n+10\Lambda_g p_0)\,\|\tilde{\mu}\|_{\varrho(-1)}.$$

For the cardinality of the set $\mathcal{K}_*(q')$, we have

$$\#\mathcal{K}_*(q') \leqslant \binom{q(n)-q'}{[\eta(q(n)-q')]}\binom{[2\eta p_0(q(n)-q')]+[\eta(q(n)-q')]}{[\eta(q(n)-q')]},$$

where the first factor on the right-hand side is an upper bound for the number of possible arrangements of integers $j\geqslant 0$ for which σ_j may be positive, and the second factor is an upper bound for the cardinality of $\sigma\in\mathcal{K}_*(q')$ when one arrangement is given. For positive numbers $\alpha,\beta>0$ and an integer $m\geqslant 1$ such that $\alpha m\geqslant 1$ and $\beta m\geqslant 1$, we have

$$\log\binom{\alpha m+\beta m}{\beta m} \leqslant \alpha m\log\left(1+\frac{\beta}{\alpha}\right)+\beta m\log\left(1+\frac{\alpha}{\beta}\right)+A_0$$

from Stirling's formula, where A_0 is an absolute constant. Hence we can obtain

$$\frac{\log\#\mathcal{K}_*(q')}{q(n)-q'} \leqslant -(1-\eta)\log(1-\eta)-\eta\log\eta+2\eta p_0\log\left(1+\frac{1}{2p_0}\right)+\eta\log(1+2p_0)+2A_0$$

for $-1\leqslant q'<q(n)$. This implies that

$$\#\mathcal{K}_*(q')\leqslant \exp(\varepsilon p_0(q(n)-q')) \quad \text{for } -1\leqslant q'<q(n),$$

provided that p_0 is larger than some constant which depends only on ε and η. Now we can conclude that

$$\begin{aligned}\|\mu|_{\mathcal{R}(n)}\circ F^{-n}\|_{\varrho} &= \|\tilde{\mu}|_{\Pi^{-1}(\mathcal{R}(n))}\circ F_*^{-n}\|_{\varrho}\\ &\leqslant \Bigg(\sum_{q'=0}^{q(n)}\sum_{\sigma\in\mathcal{K}_*(q')}\|\tilde{\mu}|_{\mathcal{D}(\sigma)}\circ F_*^{-n}\|_{\varrho}\Bigg)+\Bigg(\sum_{\sigma\in\mathcal{K}_*(-1)}\|\tilde{\mu}|_{\mathcal{D}(\sigma)}\circ F_*^{-n}\|_{\varrho}\Bigg)\\ &\leqslant \sum_{q'=0}^{q(n)}\exp(-44\varepsilon(q(n)-q')p_0+11\Lambda_g p_0+c_g)|\mu|\\ &\qquad +\exp(-44\varepsilon n+10\Lambda_g p_0)\|\mu\|_{\varrho(-1)}.\end{aligned}$$

This implies the inequality in Lemma 6.8.

7. Genericity of the transversality condition on unstable cones

In this section, we consider multiplicity of tangencies between the images of the unstable cones under iterates of mappings in $\mathcal{U}$, and investigate to what extent we can resolve the tangencies by perturbation. The goal is the proof of Theorem 3.22. The point of our argument in this section is that the dominating expansion in the unstable direction acts as uniform contraction on the angles between subspaces in the unstable cones. This enables us to control the images of the unstable cones in perturbations of mappings in $\mathcal{U}$. Notice that the content and the notation in this section is independent of those in the last section.

7.1. Reduction of Theorem 3.22: The first step

In this subsection and the next, we reduce Theorem 3.22 to more tractable propositions in two steps. For a quadruple $\chi=(\chi_c^-,\chi_c^+,\chi_u^-,\chi_u^+)$, we put

$$\chi_c^{++}:=\max\{\chi_c^+,0\},\quad \chi_c^{\Delta}:=\chi_c^+-\chi_c^- \quad\text{and}\quad \chi_u^{\Delta}:=\chi_u^+-\chi_u^-.$$

For a quadruple χ satisfying (18) and a positive number ε, let $\mathcal{S}_1(\chi,\varepsilon)$ be the set of mappings $F\in\mathcal{U}$ that satisfy

$$\limsup_{n\to\infty}\frac{1}{n}\log \mathbf{N}(\chi,\varepsilon,\varepsilon n,n;F)\geqslant \chi_c^-+\chi_u^--\chi_c^{\Delta}-\chi_u^{\Delta}-\varepsilon. \tag{74}$$

The first step of the reduction is simple. We show that we can deduce Theorem 3.22 from the following proposition:

PROPOSITION 7.1. *Suppose that $s\geqslant r+3$ and let $\mathcal{M}_s$ be the measure on $C^r(M,\mathbf{R}^2)$ introduced in Lemma 3.18. The subset $\mathcal{S}_1(\chi,\varepsilon)$ is shy with respect to the measure $\mathcal{M}_s$ for $s\geqslant r+3$, if the quadruple $\chi=(\chi_c^-,\chi_c^+,\chi_u^-,\chi_u^+)$ satisfies the conditions*

$$-2\Lambda_g<\chi_c^-<\chi_c^+<\chi_u^-<\chi_u^+<2\Lambda_g, \tag{75}$$

$$\chi_c^-<0, \tag{76}$$

$$\chi_c^{\Delta}+\chi_u^{\Delta}<\chi_c^-+\chi_u^-, \tag{77}$$

$$\chi_u^-+\chi_c^--\chi_c^{++}>\left(\frac{\chi_c^{++}+\chi_u^+}{\chi_c^-+\chi_u^--\chi_c^{\Delta}-\chi_u^{\Delta}}+1\right)(\chi_c^{\Delta}+\chi_u^{\Delta}) \tag{78}$$

and if $\varepsilon>0$ is smaller than some constant which depends only on χ and s besides the integer $r\geqslant 2$ and the objects that we fixed in §3.2.

Below we prove Theorem 3.22 assuming this proposition.

Proof of Theorem 3.22. For any point (χ_c,χ_u) in the subset given in the claim (a),

$$\{(x_c,x_u)\in\mathbf{R}^2\mid x_c+x_u>0,\ \lambda_g\leqslant x_u\leqslant\Lambda_g \text{ and } x_c\leqslant 0\},$$

we can take a quadruple $\chi=(\chi_c^-,\chi_c^+,\chi_u^-,\chi_u^+)$ satisfying the conditions (75), (76), (77) and (78) such that the rectangle $(\chi_c^-,\chi_c^+)\times(\chi_u^-,\chi_u^+)$ contains the point (χ_c,χ_u). Thus we can choose a countable collection $\mathbf{X}$ of quadruples that satisfy (75), (76), (77) and (78) such that the conditions (a) and (b) in Theorem 3.22 hold. We are going to show the condition (c) in Theorem 3.22. We fix $s\geqslant r+3$. Let $\mathbf{X}'$ be an arbitrary finite subset of $\mathbf{X}$. Then we can take a positive number $\varepsilon>0$ so small that the conclusion of Proposition 7.1 holds for all the quadruples in $\mathbf{X}'$. For each $\chi\in\mathbf{X}'$ and $n\geqslant 1$, let $\mathcal{S}_1^*(\chi,\varepsilon,n)$ be the closed subset of mappings $F\in\mathcal{U}$ that satisfy

$$\mathbf{N}(\chi,\varepsilon,\varepsilon n,n;F)\geqslant\exp((\chi_c^-+\chi_u^--\chi_c^{\Delta}-\chi_u^{\Delta}-\varepsilon)n).$$

If a mapping $F\in\mathcal{U}$ belongs to $\mathcal{S}_1(\mathbf{X}')$, or F does not satisfy the transversality condition on unstable cones for $\mathbf{X}'$, then

$$\liminf_{n\to\infty}\max\left\{\frac{\log(\mathbf{N}(\chi,\varepsilon,\varepsilon n,n;F))}{n(\chi_c^-+\chi_u^--\chi_c^{\Delta}-\chi_u^{\Delta})}\;\middle|\;\chi=(\chi_c^-,\chi_c^+,\chi_u^-,\chi_u^+)\in\mathbf{X}'\right\}\geqslant 1$$

because $\mathbf{N}(\chi,\varepsilon,k,n;F)$ is increasing with respect to ε and k. Hence we have

$$\mathcal{S}_1(\mathbf{X}')\subset\bigcup_{m>0}\bigcap_{n>m}\bigcup_{\chi\in\mathbf{X}'}\mathcal{S}_1^*(\chi,\varepsilon,n)\subset\bigcup_{\chi\in\mathbf{X}'}\mathcal{S}_1(\chi,\varepsilon).$$

From Proposition 7.1, the subset $\bigcup_{\chi\in\mathbf{X}'}\mathcal{S}_1(\chi,\varepsilon)$ is shy with respect to the measure $\mathcal{M}_s$, and hence so is $\mathcal{S}_1(\mathbf{X}')$. Further, the closed subset $\bigcap_{n>m}\bigcup_{\chi\in\mathbf{X}'}\mathcal{S}_1^*(\chi,\varepsilon,n)$ is nowhere dense, because it is shy with respect to the measure $\mathcal{M}_s$. Thus $\mathcal{S}_1(\mathbf{X}')$ is a meager subset in $\mathcal{U}$ in the sense of Baire's category argument. □

7.2. Reduction of Theorem 3.22: The second step

The second step of the reduction is rather involved. We reduce Proposition 7.1 to yet another proposition, Proposition 7.3, which will be proved in the remaining part of this section. Below we consider an integer $s\geqslant r+3$, a quadruple $\chi=\{\chi_c^-,\chi_c^+,\chi_u^-,\chi_u^+\}$ and a positive number ε. We assume that the quadruple χ satisfies the assumptions in Proposition 7.1, that is, the conditions (75), (76), (77) and (78).

In this section, we will introduce several constants that depend only on the quadruple χ and the integers $s\geqslant r\geqslant 2$ besides the objects that we fixed in §3.2. In order to distinguish such constants, we will use symbols with a subscript χ for them. Also we will use a generic symbol C_χ for large positive constants of this kind. The usage of this notation is the same as that introduced in §3.3 and §5.

The choice of the number $\varepsilon>0$ is important for our argument not only in this subsection but also in the remaining part of this section. We claim that our argument in this section is true if ε is smaller than some constant ε_χ. Below we will assume that $0<\varepsilon\leqslant\varepsilon_\chi$ and give the conditions on the choice of ε_χ in the course of the argument.

From the condition (78), we can fix a positive constant h_χ such that

$$h_\chi+1>\frac{\chi_c^{++}+\chi_u^+}{\chi_c^-+\chi_u^--\chi_c^\Delta-\chi_u^\Delta}$$

and

$$\chi_u^-+\chi_c^--\chi_c^{++}>(h_\chi+2)(\chi_c^\Delta+\chi_u^\Delta).$$

Then we fix a positive integer q_χ such that

$$q_\chi>\frac{2(\chi_u^--\chi_c^-)+\chi_c^{++}-\chi_c^-+\chi_c^\Delta+2\chi_u^\Delta}{\chi_u^-+\chi_c^--\chi_c^{++}-(h_\chi+2)(\chi_c^\Delta+\chi_u^\Delta)}.$$

Also we put

$$r_\chi=100\frac{(h_\chi+1)^2\Lambda_g^2}{\lambda_g}\geqslant 100. \tag{79}$$

Definition. For integers $0<p<n$ and a point $z\in M$, let $\mathcal{S}_1(\chi,\varepsilon,n,p,z)$ be the set of mappings $F\in\mathcal{U}$ such that there exist a subset $\{w_i\}_{i=0}^{q_\chi}$ in $F^{-p}(z)$ and subsets E_i, $0\leqslant i\leqslant q_\chi$, in $F^{-n+p}(w_i)\subset F^{-n}(z)$ that satisfy the following three conditions:

(S1) The subsets E_i for $0\leqslant i\leqslant q_\chi$ are contained in $\Lambda(\chi,\varepsilon,2(h_\chi+1)\varepsilon n,n;F)$, and

$$\#E_i=[\exp((\chi_c^-+\chi_u^--\chi_c^\Delta-\chi_u^\Delta-r_\chi\varepsilon)n)]+1.$$

(S2) For any points y and y' in the union $\bigcup_{i=0}^{q_\chi}E_i$, we have

$$\angle(DF^n(\mathbf{E}^u(y)),DF^n(\mathbf{E}^u(y')))\leqslant\exp((\chi_c^+-\chi_u^-+6\varepsilon+h_\chi(\chi_c^\Delta+\chi_u^\Delta+4\varepsilon))n).$$

(S3) For $0 \leqslant j \leqslant p$ and $0 \leqslant i, i' \leqslant q_\chi$, we have

$$F^j(\mathbf{B}(w_i, 10\exp(-r_\chi \varepsilon n)) \cap \mathbf{B}(w_{i'}, 10\exp(-r_\chi \varepsilon n)) = \varnothing,$$

except for the case where both $i=i'$ and $j=0$ hold.

For an integer $n \geqslant 1$, we consider the lattice

$$\mathbf{L}_n = \mathbf{L}(\exp((\chi_c^- - \chi_u^-)n)),$$

where $\mathbf{L}(\cdot)$ was defined in §3.1. The following lemma is the main ingredient of this subsection:

Lemma 7.2. *We have*

$$\mathcal{S}_1(\chi, \varepsilon) \subset \limsup_{n\to\infty} \bigcup_p \bigcup_{z \in \mathbf{L}_n} \mathcal{S}_1(\chi, \varepsilon, n, p, z), \tag{80}$$

where $\bigcup_p$ *indicates the union over integers* p *satisfying*

$$3h_\chi(\Lambda_g/\lambda_g)\varepsilon n \leqslant p \leqslant 3h_\chi(h_\chi+1)(\Lambda_g/\lambda_g)\varepsilon n + 1. \tag{81}$$

Proof. Let F be a mapping in $\mathcal{S}_1(\chi, \varepsilon)$. We show that there are an arbitrarily large integer n, an integer p satisfying (81) and a point $z \in \mathbf{L}_n$ such that F belongs to $\mathcal{S}_1(\chi, \varepsilon, n, p, z)$. From the definition of $\mathcal{S}_1(\chi, \varepsilon)$, there are infinitely many integers m that satisfy

$$\mathbf{N}(\chi, \varepsilon, \varepsilon m, m; F) > \exp((\chi_c^- + \chi_u^- - \chi_c^\Delta - \chi_u^\Delta - 2\varepsilon)m). \tag{82}$$

In the argument below, we consider a large integer m satisfying the condition (82). Note that, since we can take the integer m as large as we like, we may and will replace m by a larger one if it is necessary. From the definition of $\mathbf{N}(\cdot)$, there exist a point $\zeta \in M$ and a subset P in $\Lambda(\chi, \varepsilon, \varepsilon m, m; F)$ with cardinality

$$\#P > \exp((\chi_c^- + \chi_u^- - \chi_c^\Delta - \chi_u^\Delta - 2\varepsilon)m)$$

such that $F^m(P) = \{\zeta\}$ and

$$\angle(DF^m(\mathbf{E}^u(w)), DF^m(\mathbf{E}^u(w'))) \leqslant 10 H_g \exp((\chi_c^+ - \chi_u^- + 2\varepsilon)m)$$

for $w, w' \in P$. We put $p := [3h_\chi(\Lambda_g/\lambda_g)\varepsilon m] + 1$ and consider the subsets of P,

$$P_l(w) = \{w' \in P \mid F^{m-lp}(w') = F^{m-lp}(w)\}$$

for $0 \leqslant l \leqslant [m/p]$ and $w \in P$. Since $P_l(w)$ is contained in $\Lambda(\chi, \varepsilon, (m+lp)\varepsilon, m-lp; F)$ by (21), we have

$$\begin{aligned}\#P_l(w) &\leqslant \varkappa_\varepsilon \exp((\chi_u^+ + \chi_c^{++} + 7\varepsilon)(m-lp) + 6(m+lp)\varepsilon) \\ &\leqslant \exp((\chi_u^+ + \chi_c^{++} + 7\varepsilon)(m-lp) + 7(m+lp)\varepsilon)\end{aligned} \tag{83}$$

by Corollary 5.2, where the second inequality holds when m is sufficiently large. In particular, for the case $l = [m/p]$, we have

$$\#P_{[m/p]}(w) \leqslant \exp((\chi_u^+ + \chi_c^{++} + 7\varepsilon)p + 14\varepsilon m) < \exp(-[m/p]\varepsilon p)\#P,$$

where the second inequality holds if ε_χ is smaller than some constant that depends only on χ, h_χ, Λ_g and λ_g, and if we consider sufficiently large m according to the choice of ε_χ. Thus there exist integers $0 \leqslant l < [m/p]$ such that

$$\max_{w \in P} \#P_{l+1}(w) < \exp(-\varepsilon p) \max_{w \in P} \#P_l(w). \tag{84}$$

Let l_0 be the smallest integer $0 \leqslant l < [m/p]$ such that (84) holds. Then we have

$$\max_{w \in P} \#P_{l_0}(w) \geqslant \exp(-\varepsilon l_0 p)\#P.$$

Take a point $w_0 \in P$ such that $\#P_{l_0}(w_0) = \max_{w \in P} \#P_{l_0}(w)$, and put $n = m - l_0 p$, $z = F^n(w_0)$ and $E = P_{l_0}(w_0)$. Then

$$\#E = \#P_{l_0}(w_0) \geqslant \exp(-\varepsilon(m-n))\#P \geqslant \exp((\chi_c^- + \chi_u^- - \chi_c^\Delta - \chi_u^\Delta - 3\varepsilon)m).$$

Comparing this with (83) for $l = l_0$, we obtain

$$m < \frac{\chi_c^{++} + \chi_u^+}{\chi_c^- + \chi_u^- - \chi_c^\Delta - \chi_u^\Delta - 17\varepsilon} n < (h_\chi + 1)n,$$

where the second inequality follows from the choice of h_χ provided that ε_χ is smaller than some constant that depends only on χ and h_χ. Hence n and p satisfy the condition (81) and we get

$$E \subset \Lambda(\chi, \varepsilon, \varepsilon m, m; F) \subset \Lambda(\chi, \varepsilon, (m + l_0 p)\varepsilon, m - l_0 p; F) \subset \Lambda(\chi, \varepsilon, (2h_\chi + 1)\varepsilon n, n; F).$$

From (4), we can obtain, for any points w and w' in E,

$$\begin{aligned}&\angle(DF^n(\mathbf{E}^u(w)), DF^n(\mathbf{E}^u(w'))) \\ &\qquad \leqslant A_g \frac{D_* F^{m-n}(\mathbf{e}^u(F^n(w)))}{|D^* F^{m-n}(\mathbf{e}^u(F^n(w)))|} \angle(DF^m(\mathbf{E}^u(w)), DF^m(\mathbf{E}^u(w'))) \\ &\qquad \leqslant A_g \exp((-\chi_c^- + \chi_u^+)(m-n) + 2\varepsilon m)\, 10 H_g \exp((\chi_c^+ - \chi_u^- + 2\varepsilon)m) \\ &\qquad = 10 H_g A_g \exp((\chi_c^+ - \chi_u^- + 4\varepsilon)n + (\Delta\chi_c + \Delta\chi_u + 4\varepsilon)(m-n)) \\ &\qquad \leqslant \exp((\chi_c^+ - \chi_u^- + 5\varepsilon + h_\chi(\chi_c^\Delta + \chi_u^\Delta + 4\varepsilon))n),\end{aligned}$$

provided that m is sufficiently large.

Let us consider the subset $\{w_i\}_{i=1}^{i_0} \subset F^{-p}(z)$ of all points $w \in F^{-p}(z)$ such that $F^{n-p}(w) \cap E \neq \emptyset$. By (19) and (21), it is contained in $\Lambda(\chi, \varepsilon, \varepsilon(m+pl_0), p; F)$. Corollary 5.2 gives the following estimate for its cardinality i_0:

$$i_0 \leqslant \varkappa_\varepsilon \exp(5\Lambda_g p + 6\varepsilon(m+pl_0)) \leqslant \varkappa_\varepsilon \exp(5\Lambda_g p + 12\varepsilon m).$$

We put $E_i = \{y \in E \mid F^{n-p}(y) = w_i\}$ for $1 \leqslant i \leqslant i_0$, so that $E = \bigcup_{i=1}^{i_0} E_i$.

By changing the index i, we assume that the cardinality of the subset E_i is decreasing with respect to i. Let i_1 be the smallest positive integer such that

$$\sum_{i=1}^{i_1} \# E_i > \frac{1}{2} \sum_{i=1}^{i_0} \# E_i = \frac{\# E}{2}.$$

Then we have $\# E_{i_1} \cdot (i_0 - i_1 + 1) \geqslant \sum_{i=i_1}^{i_0} \# E_i \geqslant \frac{1}{2} \# E$ and hence

$$\# E_i \geqslant \# E_{i_1} \geqslant \frac{\# E}{2 i_0} > \exp((\chi_c^- + \chi_u^- - \chi_c^\Delta - \chi_u^\Delta - r_\chi \varepsilon) n) \quad \text{for } 1 \leqslant i \leqslant i_1,$$

where the last inequality follows from the definitions of p and r_χ provided that m is sufficiently large. We also have

$$i_1 \geqslant \frac{\sum_{i=1}^{i_1} \# E_i}{\# E_1} \geqslant \frac{\# E}{2 \# E_1} \geqslant \frac{\exp \varepsilon p}{2}$$

from the condition (84) for $l = l_0$.

Notice that the point z that we took above may not be contained in $\mathbf{L}_n$, while we would like it to be. So we want to shift it to the closest point in $\mathbf{L}_n$. The distance from the point z to the closest point in $\mathbf{L}_n$ is bounded by $\exp((\chi_c^- - \chi_u^-) n)$, and hence by $\varrho_\varepsilon \exp((\chi_c^- - 5\varepsilon - 3(2h_\chi + 1)\varepsilon) n)$, provided that ε_χ is smaller than some constant which depends only on χ and that we took sufficiently large m. Thereby, by virtue of Lemma 5.1, we can move the points w_i and those in E_i accordingly so that the relations $F^p(w_i) = z$ and $F^{n-p}(E_i) = \{w_i\}$ are preserved. Henceforth, we consider the points $z \in \mathbf{L}_n$, w_i and the subsets E_i thus obtained. Lemma 5.1 guarantees that the subsets E_i are contained in $\Lambda(\chi, \varepsilon, 2(h_\chi + 1)\varepsilon n, n; F)$ and that

$$\begin{aligned}\angle(DF^n(\mathbf{E}^u(w)), DF^n(\mathbf{E}^u(w'))) &\leqslant \exp((\chi_c^+ - \chi_u^- + 5\varepsilon + h_\chi(\chi_c^\Delta + \chi_u^\Delta + 4\varepsilon)) n) \\ &\quad + 2\varkappa_\varepsilon \exp((\chi_c^- - \chi_u^- + (4h_\chi + 2)\varepsilon) n) \\ &< \exp((\chi_c^+ - \chi_u^- + 6\varepsilon + h_\chi(\chi_c^\Delta + \chi_u^\Delta + 4\varepsilon)) n)\end{aligned}$$

for any points $w, w' \in \bigcup_{i=1}^{i_1} E_i$, provided that ε_χ is smaller than some constant which depends only on χ and that m is sufficiently large. Up to this point, we have found an arbitrarily large integer n, an integer p, points z, w_i, $1 \leqslant i \leqslant i_1$, and subsets E_i, $1 \leqslant i \leqslant i_1$, that satisfy the conditions (81), (S1) and (S2). It remains to choose $q_\chi+1$ points among w_i, $1 \leqslant i \leqslant i_1$, so that the condition (S3) holds.

Put $W=\{w_i \mid 1 \leqslant i \leqslant i_1\}$ and $\delta=40p\exp(2\Lambda_g p-r_\chi \varepsilon n)$. Note that the points w_i belong to $\Lambda(\chi, \varepsilon, 2(h_\chi+1)\varepsilon n, p; F)$ by (19). We can check that

$$2\delta < \varkappa_g^{-1} \varrho_\varepsilon \exp((-\chi_u^+ + \chi_c^- - 5\varepsilon)p - 8(h_\chi+1)\varepsilon n)$$

by using the definition of p and r_χ, and the condition (81), provided that m is large enough. Thus F^p is a diffeomorphism on the 2δ-neighborhood of each point in W from Lemma 5.1 (v). This implies that the distances between the points in $W \subset F^{-p}(z)$ are not less than 2δ. Let $L \subset W$ be the set of points in W that are within distance δ to either of the points $F^j(z)$, $0 \leqslant j < p$. Then we obviously have $\# L \leqslant p$.

Consider a sequence $J=\{j_\nu\}_{\nu=0}^{\nu_0}$ of integers such that $1 \leqslant j_\nu \leqslant p$ for $0 \leqslant \nu \leqslant \nu_0$. The sum of the integers in J is denoted by $|J| := \sum_{\nu=0}^{\nu_0} j_\nu$. For $x, x' \in W \setminus L$, we write $x \succ_J x'$ if there is a sequence of points $x_0=x, x_1, \ldots, x_{\nu_0+1}=x'$ in $W \setminus L$ such that

$$F^{j_\nu}(\mathbf{B}(x_\nu, 10\exp(-r_\chi \varepsilon n))) \cap \mathbf{B}(x_{\nu+1}, 10\exp(-r_\chi \varepsilon n)) \neq \varnothing \quad \text{for } 0 \leqslant \nu \leqslant \nu_0.$$

From the definition of δ above, it is easy to see that we have $d(F^{|J|}(x), x') < \delta$ if $x \succ_J x'$ for some J with $|J| \leqslant 2p$. Hence, given a point $x \in W \setminus L$ and an integer $1 \leqslant i \leqslant 2p$, there is at most one point x' in $W \setminus L$ that satisfies $x \succ_J x'$ for some sequence J with $|J|=i$.

Actually, the relation $x \succ_J x'$ holds for some points x and x' in $W \setminus L$ only if $|J|<p$. In fact, otherwise, there should be a sequence J with $p \leqslant |J| < 2p$ and points x and x' in $W \setminus L$ such that $x \succ_J x'$, and hence $d(F^{|J|-p}(z), x') = d(F^{|J|}(x), x') < \delta$. But, since $0 \leqslant |J|-p < p$, this contradicts the definition of L.

The relation $x \succ_J x'$ never holds if $x=x'$. In fact, if $x \succ_J x$ for some J, the relation $x \succ_{J^i} x$ should hold for any $i \geqslant 1$, where J^i is the iteration of J, i times. But this obviously contradicts the fact proved in the preceding paragraph.

We write $x \succ x'$ for $x, x' \in W \setminus L$ if either $x=x'$, or $x \succ_J x'$ for some sequence $J=\{j_\nu\}_{\nu=0}^{\nu_0}$ satisfying $1 \leqslant j_\nu \leqslant p$. From the argument above, this relation is a partial order on the set $W \setminus L$, and, for each $x \in W \setminus L$, there exist at most p points x' in $W \setminus L$ such that $x \succ x'$. Let $W_{\max}$ be the set of the maximal elements in $W \setminus L$ with respect to the partial order $\succ$. Then we have

$$\# W_{\max} \geqslant \frac{\#(W \setminus L)}{p} \geqslant \frac{\left[\frac{1}{2}\exp \varepsilon p\right] - p}{p} \geqslant q_\chi + 1,$$

provided that m is large enough. Take $q_\chi+1$ points $\{w_i\}_{i=0}^{q_\chi}$ from $W_{\max}$. Then the condition (S3) holds for them. We have completed the proof of Lemma 7.2. □

Using Lemma 7.2, we can deduce Proposition 7.1 from the following proposition:

PROPOSITION 7.3. *Let $s\geqslant r+3$. Suppose that a quadruple χ satisfies the conditions* (75), (76), (77) *and* (78), *and that a positive number ε satisfies $0<\varepsilon\leqslant\varepsilon_\chi$. Then, for any $d>0$ and any mapping G in $C^r(M,\mathbf{T})$, there exists an integer n_0 such that*

$$\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_1(\chi,\varepsilon,n,p,z))\cap\mathbf{D}^{s-3}(d))<\exp((2\chi_c^- -2\chi_u^- -\varepsilon)n) \tag{85}$$

for $n\geqslant n_0$, $z\in\mathbf{L}_n$ and $0<p<n$ satisfying the condition (81).

Remark. Φ_G and $\mathbf{D}^{s-3}(d)$ above are defined by (2) and (25), respectively.

In fact, since we have $\#L_n=([\exp((-\chi_c^- +\chi_u^-)n)]+1)^2$, it follows from Proposition 7.3 and (80) that

$$\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_1(\chi,\varepsilon))\cap\mathbf{D}^{s-3}(d))=0 \quad \text{for any } d>0 \text{ and } G\in C^r(M,\mathbf{T}).$$

Since the measure $\mathcal{M}_s$ is supported on $C^{s-3}(M,\mathbf{R}^2)=\bigcup_{d>0}\mathbf{D}^{s-3}(d)$, this implies that the subset $\mathcal{S}_1(\chi,\varepsilon)$ is shy with respect to the measure $\mathcal{M}_s$.

7.3. Perturbations

In this subsection, we introduce some families of mappings and give estimates on the variations of the images of the unstable subspaces $\mathbf{E}^u(z)$ under iterates of the mappings in the families. Henceforth, in this subsection and the next, we consider the situation in Proposition 7.3: Let $s\geqslant r+3$, let χ be a quadruple that satisfies the conditions (75), (76), (77) and (78), and let ε be a positive number that satisfies $0<\varepsilon\leqslant\varepsilon_\chi$.

Fix a C^∞-function $\psi:\mathbf{R}^2\to\mathbf{R}$ such that $\|\psi\|_{C^1}\leqslant 1$ and

$$\psi(w)=\begin{cases} x, & \text{if } \|w\|\leqslant\frac{1}{10},\\ 0, & \text{if } \|w\|\geqslant 1,\end{cases}$$

for $w=(x,y)\in\mathbf{R}^2$. For each point $z\in M$, we consider an isometric embedding

$$\varphi_z:\{w\in\mathbf{R}^2 \mid \|w\|<\tfrac{1}{5}\}\longrightarrow\mathbf{T}$$

that carries the origin to z and the x-axis $\mathbf{R}\times\{0\}$ to $\mathbf{E}^u(z)$. For $n\geqslant 1$, we put

$$\delta_n=\exp(-r_\chi\varepsilon n).$$

Recall that we took the subset $\mathcal{U}$ of mappings as a neighborhood of a C^r-mapping $F_\sharp$ in §3.2. For an integer $n \geqslant 1$ and a point $z \in M$, we define the C^∞-mapping $\psi_{n,z}: M \to \mathbf{R}^2$ by

$$\psi_{n,z}(w) := \begin{cases} \delta_n^{s+3} \psi(\varphi_z^{-1}(w)/\delta_n)\, \mathbf{e}^c(F_\sharp(z)), & \text{if } d(w,z) < \delta_n, \\ 0, & \text{otherwise,} \end{cases}$$

where $\mathbf{e}^c(\,\cdot\,)$ is either of the two unit vectors in the central subspace $\mathbf{E}^c(\,\cdot\,)$. Note that, for any mapping $F \in \mathcal{U}$, the parallel translation of the vector $\mathbf{e}^c(F_\sharp(z))$ to $F(z)$ is contained in $\mathbf{S}^c(F(z))$ from the choice of the constant ϱ_g in §3.2.

Remark. Notice that the definition of $\psi_{n,z}(w)$ does not depend on $F \in \mathcal{U}$.

Let n and p be positive integers that satisfy the condition (81), $S = \{x_i\}_{i=0}^{q_\chi}$ an ordered subset of the lattice $\mathbf{L}(\frac{1}{40}\delta_n)$, and F a mapping in $\mathcal{U}$. The family of mappings that we are going to consider is

$$F_{\mathbf{t}}(w) = F(w) + \sum_{i=1}^{q_\chi} t_i \psi_{n,x_i}(w): M \longrightarrow \mathbf{T},$$

where $\mathbf{t} = \{t_i\}_{i=1}^{q_\chi} \in \mathbf{R}^{q_\chi}$ is the parameter that ranges over the region

$$R = \big\{ \mathbf{t} = \{t_i\}_{i=1}^{q_\chi} \in \mathbf{R}^{q_\chi} \;\big|\; |t_i| \leqslant \exp(\chi_c^- n) \big\}.$$

For this family, we have

$$d_{C^l}(F_{\mathbf{t}}, F) \leqslant C_g q_\chi \delta_n^{s-l+3} \|\mathbf{t}\| \cdot \|\psi\|_{C^l} \quad \text{for } \mathbf{t} \in R \text{ and } 0 \leqslant l \leqslant s. \tag{86}$$

From this inequality in the case $l=0$, we obtain

$$d_{C^0}(F_{\mathbf{t}}^j, F^j) \leqslant p \exp(\Lambda_g p)\, C_g q_\chi \delta_n^{s+3} \exp(\chi_c^- n) < \delta_n \tag{87}$$

for $0 \leqslant j \leqslant p$ and $\mathbf{t} \in R$, where the second inequality follows from the condition (81) and the definition of r_χ provided that n is larger than some constant N_ε. (Recall the notation introduced in §5.)

Let us use the notation ∂_i for the partial differentiation with respect to the parameter t_i. We have

$$\|\partial_i F_{\mathbf{t}}(w)\| \leqslant \delta_n^{s+3} \tag{88}$$

and

$$\|\partial_i (DF_{\mathbf{t}})(\mathbf{v})\| \leqslant C_g \delta_n^{s+2} \|\mathbf{v}\| \tag{89}$$

for any $w \in M$, $\mathbf{v} \in \mathbf{S}^u(w)$ and $\mathbf{t} \in R$. If $d(w, x_i) < \frac{1}{10}\delta_n$ in addition, we also have

$$|\mathbf{v}^*(\partial_i (DF_{\mathbf{t}}(\mathbf{v})))| \geqslant C_g^{-1} \delta_n^{s+2} \|\mathbf{v}\| \tag{90}$$

for any $w \in M$, $\mathbf{v} \in \mathbf{S}^u(w)$ and $\mathbf{t} \in R$, where $\mathbf{v}^*$ is the unit cotangent vector at $F_{\mathbf{t}}(w)$ that is normal to $DF_{\mathbf{t}}(\mathbf{v})$.

In the following argument, we assume that

$$F^j(\mathbf{B}(x_i, 2\delta_n)) \cap \mathbf{B}(x_{i'}, 2\delta_n) = \varnothing \tag{91}$$

for $0 \leqslant i, i' \leqslant q_\chi$ and $0 \leqslant j \leqslant p$, except for the case where both $i=i'$ and $j=0$ hold. Note that (91) and the estimate (87) imply that

$$F_{\mathbf{t}}^j(\mathbf{B}(x_i, \delta_n)) \cap \mathbf{B}(x_{i'}, \delta_n) = \varnothing \quad \text{for } \mathbf{t} \in R. \tag{92}$$

Consider a point $z \in M$ and families of points $y_i(\mathbf{t}) \in M$, $0 \leqslant i \leqslant q_\chi$, parameterized by $\mathbf{t} \in R$ continuously. Suppose that

(Y1) $F_{\mathbf{t}}^n(y_i(\mathbf{t})) = z$;

(Y2) $y_i(\mathbf{t}) \in \Lambda(\chi, \varepsilon, (2h_\chi+3)\varepsilon n, n; F_{\mathbf{t}})$;

(Y3) $d(F_{\mathbf{t}}^{n-p}(y_i(\mathbf{t})), x_i) < \frac{1}{10}\delta_n$

for $0 \leqslant i \leqslant q_\chi$ and $\mathbf{t} \in R$. Let us put

$$A_i(\mathbf{t}) = \delta_n^{s+2} \frac{|D^* F_{\mathbf{t}}^{p-1}(DF_{\mathbf{t}}^{n-p+1}(\mathbf{e}^u(y_i(\mathbf{t}))))|}{D_* F_{\mathbf{t}}^{p-1}(DF_{\mathbf{t}}^{n-p+1}(\mathbf{e}^u(y_i(\mathbf{t}))))}$$

for $1 \leqslant i \leqslant q_\chi$, where $\mathbf{e}^u(z)$ is either of the two unit tangent vectors in $\mathbf{E}^u(z)$. Then we can show the following estimates on the motion of the subspace $DF_{\mathbf{t}}^n(\mathbf{E}^u(y_i(\mathbf{t})))$ as the parameter $\mathbf{t}$ moves:

LEMMA 7.4. *Let the constant N_ε be larger if necessary. If $n \geqslant N_\varepsilon$, we have*

$$C_g^{-1} A_i(\mathbf{t}) \leqslant |\partial_i \angle(DF_{\mathbf{t}}^n(\mathbf{E}^u(y_i(\mathbf{t}))), \mathbf{E}^u(z))| \leqslant C_g A_i(\mathbf{t})$$

for $1 \leqslant i \leqslant q_\chi$, and also

$$|\partial_j \angle(DF_{\mathbf{t}}^n(\mathbf{E}^u(y_i(\mathbf{t}))), \mathbf{E}^u(z))| \leqslant C_g \exp(-\lambda_g p) A_i(\mathbf{t})$$

for $0 \leqslant i \leqslant q_\chi$ and $1 \leqslant j \leqslant q_\chi$, provided that $i \neq j$.

Proof. Let $1 \leqslant i \leqslant q_\chi$ and $0 \leqslant j \leqslant q_\chi$. For $0 \leqslant m \leqslant n$, let $\mathbf{e}_m$ be the unit tangent vector in the direction of $DF_{\mathbf{t}}^m(\mathbf{e}^u(y_i(\mathbf{t})))$, and denote by $\mathbf{e}_m^*$ the unit cotangent vector that is normal to $\mathbf{e}_m$. We can choose the orientation of the cotangent vectors $\mathbf{e}_m^*$ so that $(DF^{n-m})^*(\mathbf{e}_n^*) = D^* F^{n-m}(\mathbf{e}_m)\mathbf{e}_m^*$. Also we put $z_m = F_{\mathbf{t}}^m(y_i(\mathbf{t}))$ for simplicity. Notice that $\mathbf{e}_m$, $\mathbf{e}_m^*$ and z_m depend on the parameter $\mathbf{t}$.

We first give some simple consequences of the conditions (Y1) and (Y3). By (92) and the condition (Y3), the point z_m is not contained in $\mathbf{B}(x_j, \delta_n)$ for $n-2p < m < n$, except for the case where both $m=n-p$ and $j=i$ hold. In particular, the points z_m for $n-p<m<n$ are not contained in $\bigcup_{l=0}^{q_x}\mathbf{B}(x_l,\delta_n)$. So the condition (Y1) implies that the point z_m for $n-p<m\leqslant n$ does not depend on the parameter $\mathbf{t}$. For $0\leqslant m\leqslant n-p$, differentiation of both sides of the identity $F_{\mathbf{t}}^{n-p+1-m}(z_m)\equiv z_{n-p+1}$ gives

$$(DF_{\mathbf{t}}^{n-p+1-m})_{z_m}(\partial_j z_m)+\sum_{l=m+1}^{n-p+1}(DF_{\mathbf{t}}^{n-p+1-l})_{z_l}((\partial_j F_{\mathbf{t}})(z_{l-1}))=0.$$

Applying $(DF_{\mathbf{t}}^{n-p+1-m})_{z_m}^{-1}$ to both sides of this identity and using (88) and (7), we obtain

$$|\partial_j z_m|\leqslant\sum_{l=m+1}^{n-p+1}C_g\|((DF_{\mathbf{t}}^{l-m})_{z_m})^{-1}\|\,\delta_n^{s+3}\leqslant\sum_{l=m+1}^{n-p+1}\frac{C_g\delta_n^{s+3}}{|D^*F_{\mathbf{t}}^{l-m}(\mathbf{e}_m)|}. \tag{93}$$

Now we are going to estimate

$$\partial_j\angle(DF_{\mathbf{t}}^n(\mathbf{E}^u(y_i(\mathbf{t}))),\mathbf{E}^u(z))=\partial_j\angle(\mathbf{e}_n,\mathbf{E}^u(z))=\frac{\mathbf{e}_n^*(\partial_j(DF_{\mathbf{t}}^n(\mathbf{e}_0)))}{D_*F_{\mathbf{t}}^n(\mathbf{e}_0)}.$$

Differentiating both sides of

$$DF_{\mathbf{t}}^n(\mathbf{e}_0)=(DF_{\mathbf{t}})_{z_{n-1}}\circ(DF_{\mathbf{t}})_{z_{n-2}}\circ\ldots\circ(DF_{\mathbf{t}})_{z_0}(\mathbf{e}_0)$$

and using the relation $DF_{\mathbf{t}}^m(\mathbf{e}_0)=D_*F_{\mathbf{t}}^m(\mathbf{e}_0)\,\mathbf{e}_m$, we can obtain

$$\begin{aligned}\partial_j(DF_{\mathbf{t}}^n(\mathbf{e}_0))&=\sum_{m=0}^{n-1}(DF_{\mathbf{t}}^{n-m-1})_{z_{m+1}}((\partial_j(DF_{\mathbf{t}})_{z_m})(\mathbf{e}_m))\,D_*F_{\mathbf{t}}^m(\mathbf{e}_0)\\&\quad+\sum_{m=0}^{n-1}(DF_{\mathbf{t}}^{n-m-1})_{z_{m+1}}(D^2F_{\mathbf{t}}(\mathbf{e}_m,\partial_j z_m))\,D_*F_{\mathbf{t}}^m(\mathbf{e}_0)\\&\quad+(DF_{\mathbf{t}}^n)_{z_0}(D\mathbf{e}^u(\partial_j z_0)).\end{aligned}$$

From this and the relation $(DF^{n-m})^*(\mathbf{e}_n^*)=D^*F^{n-m}(\mathbf{e}_m)\,\mathbf{e}_m^*$, it follows that

$$\begin{aligned}\mathbf{e}_n^*(\partial_j(DF_{\mathbf{t}}^n(\mathbf{e}_0)))&=\sum_{m=0}^{n-1}D^*F_{\mathbf{t}}^{n-m-1}(\mathbf{e}_{m+1})\,\mathbf{e}_{m+1}^*((\partial_j(DF_{\mathbf{t}})_{z_m})(\mathbf{e}_m))\,D_*F_{\mathbf{t}}^m(\mathbf{e}_0)\\&\quad+\sum_{m=0}^{n-1}D^*F_{\mathbf{t}}^{n-m-1}(\mathbf{e}_{m+1})\,\mathbf{e}_{m+1}^*(D^2F_{\mathbf{t}}(\mathbf{e}_m,\partial_j z_m))\,D_*F_{\mathbf{t}}^m(\mathbf{e}_0)\\&\quad+D^*F_{\mathbf{t}}^n(\mathbf{e}_0)\,\mathbf{e}_0^*(D\mathbf{e}^u(\partial_j z_0)).\end{aligned}$$

Note that $(\partial_j(DF_{\mathbf{t}}))_{z_m}=0$ for $n-2p<m<n$, except for the case $m=n-p$, and that $\partial_j z_m=0$ for $n-p<m\leqslant n$ as we noted above. Thus we obtain

$$\begin{aligned}
\frac{\mathbf{e}_n^*(\partial_j(DF_{\mathbf{t}}^n(\mathbf{e}_0)))}{D_*F_{\mathbf{t}}^n(\mathbf{e}_0)} - \frac{D^*F_{\mathbf{t}}^{p-1}(\mathbf{e}_{n-p+1})}{D_*F_{\mathbf{t}}^{p-1}(\mathbf{e}_{n-p+1})}\,\frac{\mathbf{e}_{n-p+1}^*((\partial_j(DF_{\mathbf{t}})_{z_{n-p}})(\mathbf{e}_{n-p}))}{D_*F_{\mathbf{t}}(\mathbf{e}_{n-p})}
&= \sum_{m=0}^{n-2p}\frac{D^*F_{\mathbf{t}}^{n-m-1}(\mathbf{e}_{m+1})}{D_*F_{\mathbf{t}}^{n-m-1}(\mathbf{e}_{m+1})}\,\frac{\mathbf{e}_{m+1}^*((\partial_j(DF_{\mathbf{t}})_{z_m})(\mathbf{e}_m))}{D_*F_{\mathbf{t}}(\mathbf{e}_m)}\\
&\quad+\sum_{m=0}^{n-p}\frac{D^*F_{\mathbf{t}}^{n-m-1}(\mathbf{e}_{m+1})}{D_*F_{\mathbf{t}}^{n-m-1}(\mathbf{e}_{m+1})}\,\frac{\mathbf{e}_{m+1}^*(D^2F_{\mathbf{t}}(\mathbf{e}_m,\partial_j z_m))}{D_*F_{\mathbf{t}}(\mathbf{e}_m)}\\
&\quad+\frac{D^*F_{\mathbf{t}}^n(\mathbf{e}_0)}{D_*F_{\mathbf{t}}^n(\mathbf{e}_0)}\,\mathbf{e}_0^*(D\mathbf{e}^u(\partial_j z_0)).
\end{aligned}\tag{94}$$

From (89), the first sum on the right-hand side is bounded in absolute value by

$$C_g\delta_n^{s+2}\frac{|D^*F_{\mathbf{t}}^{p-1}(\mathbf{e}_{n-p+1}^u)|}{D_*F_{\mathbf{t}}^{p-1}(\mathbf{e}_{n-p+1}^u)}\sum_{m=0}^{n-2p}\exp(-\lambda_g(n-p+1-m+2c_g))\leqslant C_gA_i(\mathbf{t})\exp(-\lambda_g p).$$

By the estimate (93) on $\partial_j z_m$ and the condition (Y2), the second sum on the right-hand side is bounded in absolute value by

$$\begin{aligned}
&C_g\sum_{m=0}^{n-p}\sum_{l=m+1}^{n-p+1}\frac{|D^*F_{\mathbf{t}}^{n-m-1}(\mathbf{e}_{m+1})|}{D_*F_{\mathbf{t}}^{n-m}(\mathbf{e}_m)}\,\frac{\delta_n^{s+3}}{|D^*F_{\mathbf{t}}^{l-m}(\mathbf{e}_m)|}\\
&\quad=C_g\sum_{m=0}^{n-p}\sum_{l=m+1}^{n-p+1}\frac{|D^*F_{\mathbf{t}}^{n-l}(\mathbf{e}_l)|}{D_*F_{\mathbf{t}}^{n-l}(\mathbf{e}_l)}\,\frac{\delta_n^{s+3}}{D_*F_{\mathbf{t}}^{l-m}(\mathbf{e}_m)|D^*F_{\mathbf{t}}(\mathbf{e}_m)|}\\
&\quad<C_g\delta_n^{s+3}\frac{|D^*F_{\mathbf{t}}^{p-1}(\mathbf{e}_{n-p+1})|}{D_*F_{\mathbf{t}}^{p-1}(\mathbf{e}_{n-p+1})}\sum_{m=0}^{n-p}\sum_{l=m+1}^{n-p+1}\frac{\exp(-\lambda_g(n-p+1-m+2c_g))}{\exp(-(2h_\chi+4)\varepsilon n)}\\
&\quad<C_gA_i(\mathbf{t})\delta_n\exp((2h_\chi+4)\varepsilon n)\\
&\quad<C_gA_i(\mathbf{t})\exp(-\lambda_g p),
\end{aligned}$$

where the last inequality follows from the definition of the constant r_χ and the condition (81) on p. Similarly, we can show that the last term on the right-hand side is bounded by

$$C_g\sum_{l=1}^{n-p+1}\frac{|D^*F_{\mathbf{t}}^n(\mathbf{e}_0)|}{D_*F_{\mathbf{t}}^n(\mathbf{e}_0)}\,\frac{\delta_n^{s+3}}{|D^*F_{\mathbf{t}}^l(\mathbf{e}_0)|}<C_gA_i(\mathbf{t})\exp(-\lambda_g p).$$

From (89) and (90), we have

$$C_g^{-1}\delta_n^{s+2}<|\mathbf{e}_{n-p+1}^*(\partial_j DF_{\mathbf{t}}(\mathbf{e}_{n-p}))|<C_g\delta_n^{s+2}\quad\text{if } j=i,$$

and $\partial_i DF_{\mathbf{t}}(\mathbf{e}_{n-p})\equiv 0$ otherwise. Using these estimates in (94), we can conclude the lemma, by taking the constant N_ε larger if necessary. □

Consider the mapping $\Psi: R\to\mathbf{R}^{q_\chi}$ defined by

$$\Psi(\mathbf{t}) = \{\angle(DF_{\mathbf{t}}^n(\mathbf{E}^u(y_i(\mathbf{t}))), DF_{\mathbf{t}}^n(\mathbf{E}^u(y_0(\mathbf{t}))))\}_{i=1}^{q_\chi}. \tag{95}$$

As a consequence of Lemma 7.4, we have the following corollary, where we take the constant N_ε still larger if necessary:

COROLLARY 7.5. *The mapping Ψ is injective and there is a constant B_χ such that*

$$|\det D\Psi(\mathbf{t})| > \exp(-B_\chi \varepsilon n) \quad \textit{for } \mathbf{t}\in R,$$

provided that $n\geqslant N_\varepsilon$.

Proof. Let $D\Psi(\mathbf{t})_{ij}$ be the (i,j)-entry of the representation matrix of $D\Psi(\mathbf{t})$ with respect to the standard coordinate on $\mathbf{R}^{q_\chi}$. Lemma 7.4 tells us that the diagonal entries satisfy

$$C_g^{-1}A_i(\mathbf{t}) < |D\Psi(\mathbf{t})_{ii}| < C_g A_i(\mathbf{t}),$$

while the off-diagonal entries satisfy

$$|D\Psi(\mathbf{t})_{ij}| < C_g \exp(-\lambda_g p) A_j(\mathbf{t}), \quad j\neq i.$$

These facts imply that Ψ is injective on R and that $|\det D\Psi(\mathbf{t})|$ is bounded from below by $\prod_{i=1}^{q_\chi} C_g A_i(\mathbf{t})$, provided that n is larger than some constant C_χ. Therefore we have

$$|\det D\Psi(\mathbf{t})| > (C_g \exp((\chi_c^- - \chi_u^+)p - (4h_\chi + 6 + (s+2)r_\chi)\varepsilon n))^{q_\chi}$$

from the condition (Y2). Using the condition (81), we obtain the corollary. □

7.4. The proof of Proposition 7.3

In this subsection, we complete the proof of Theorem 3.22 by proving Proposition 7.3. Let G be a mapping in $C^r(M,\mathbf{T})$ and $d>0$ a positive number. We consider a large integer $n>N_\varepsilon$, an integer p satisfying the condition (81), and a point z in the lattice $\mathbf{L}_n$. We put $\delta_n=\exp(-r_\chi\varepsilon n)$ as in the last subsection. Consider an ordered subset $S=\{x_i\}_{i=0}^{q_\chi}$ in the lattice $\mathbf{L}(\frac{1}{40}\delta_n)$. Let $\mathcal{S}_1(\chi,\varepsilon,n,p,z;S)$ be the set of mappings F in $\mathcal{S}_1(\chi,\varepsilon,n,p,z)$ such that the subset $\{w_i\}_{i=0}^{q_\chi}$ in the definition can be taken so that

(S4) $d(w_i,x_i)<\frac{1}{20}\delta_n$ for $0\leqslant i\leqslant q_\chi$.

The subset $\mathcal{S}_1(\chi,\varepsilon,n,p,z)$ is contained in the union of $\mathcal{S}_1(\chi,\varepsilon,n,p,z;S)$ over all ordered subsets $S=\{x_i\}_{i=0}^{q_\chi}$ of the lattice $\mathbf{L}(\frac{1}{40}\delta_n)$. And the number of such ordered sets S is bounded by $(40\delta_n^{-1}+1)^{2(q_\chi+1)}$. Therefore, in order to prove the inequality in Proposition 7.3, it is enough to show that

$$\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_1(\chi,\varepsilon,n,p,z;S))\cap\mathbf{D}^{s-3}(d))<\exp((2(\chi_c^- -\chi_u^-)-2r_\chi(q_\chi+2)\varepsilon)n) \tag{96}$$

for sufficiently large n.

Take an arbitrary mapping F in $\mathcal{S}_1(\chi,\varepsilon,n,p,z;S)$ and consider the family of mappings $F_{\mathbf{t}}$ defined for the ordered subset S in the last subsection. Note that the conditions (91) and (92) follow from the conditions (S3) and (S4). Let $\mathcal{Y}$ be the set of continuous mappings

$$\mathbf{y}: R\longrightarrow M\times M\times\ldots\times M,\quad \mathbf{y}(\mathbf{t})=\{y_i(\mathbf{t})\}_{i=0}^{q_\chi},$$

that satisfy the conditions (Y1), (Y2) and (Y3) in the last subsection. A family $\mathbf{y}(\mathbf{t})$ in $\mathcal{Y}$ is uniquely determined once $\mathbf{y}(0)$ is given because of the conditions (Y1) and (Y2). Thus we have

$$\begin{aligned}\#\mathcal{Y}&\leqslant(\#(\Lambda(\chi,\varepsilon,(2h_\chi+3)\varepsilon n,n;F)\cap F^{-n}(z)))^{q_\chi+1}\\&\leqslant\varkappa_\varepsilon\exp((\chi_u^+ +\chi_c^{++}+7\varepsilon+6(2h_\chi+3)\varepsilon)(q_\chi+1)n)\\&\leqslant\exp((\chi_u^+ +\chi_c^{++})(q_\chi+1)n+C_\chi\varepsilon n)\end{aligned}$$

for sufficiently large n, by Corollary 5.2 and the condition (Y2).

For a family $\mathbf{y}\in\mathcal{Y}$, let $Z(\mathbf{y})$ be the set of parameters $\mathbf{t}\in R$ such that

$$\angle(DF_{\mathbf{t}}^n(\mathbf{E}^u(y_i(\mathbf{t}))),DF_{\mathbf{t}}^n(\mathbf{E}^u(y_0(\mathbf{t}))))\leqslant\exp((\chi_c^+ -\chi_u^- +6\varepsilon+h_\chi(\chi_c^\Delta+\chi_u^\Delta+4\varepsilon))n)$$

for all $1\leqslant i\leqslant q_\chi$. Corollary 7.5 implies that we have

$$\mathbf{m}(Z(\mathbf{y}))\leqslant\exp((\chi_c^+ -\chi_u^- +h_\chi(\chi_c^\Delta+\chi_u^\Delta))q_\chi n+C_\chi\varepsilon n),$$

provided that $n\geqslant N_\varepsilon$.

Suppose that $F_{\mathbf{s}}$ belongs to $\mathcal{S}_1(\chi,\varepsilon,n,p,z;S)$ for a parameter $\mathbf{s}\in R$. Then there are points $w_i\in F_{\mathbf{s}}^{-p}(z)$, $0\leqslant i\leqslant q_\chi$, and subsets $E_i\subset F_{\mathbf{s}}^{-(n-p)}(w_i)$, $0\leqslant i\leqslant q_\chi$, which satisfy the conditions (S1)–(S4) with F replaced by $F_{\mathbf{s}}$. Consider a combination $\{y_i\}_{i=0}^{q_\chi}$ of points such that $y_i\in E_i$ for $0\leqslant i\leqslant q_\chi$. From (86), we can check that

$$d_{C^1}(F_{\mathbf{t}},F_{\mathbf{s}})<\varrho_\varepsilon\exp((\chi_c^- -5\varepsilon)n-3\cdot 2(h_\chi+1)\varepsilon n)\quad\text{for any }\mathbf{t}\in R,$$

provided that n is sufficiently large. Thus, by the condition (S1) and Lemma 5.1, we can check that there exists a unique element $\mathbf{y}(\mathbf{t})=\{y_i(\mathbf{y})\}_{i=0}^{q_\chi}$ in $\mathcal{Y}$ such that $y_i(\mathbf{s})=y_i$ for

$0 \leqslant i \leqslant q_\chi$. The condition (S2) implies that $\mathbf{s}$ belongs to the subset $Z(\mathbf{y})$. Therefore, if $F_\mathbf{s}$ belongs to $\mathcal{S}_1(\chi,\varepsilon,n,p,z;S)$, the parameter $\mathbf{s}$ belongs to the subset $Z(\mathbf{y})$ for at least

$$\prod_{i=0}^{q_\chi} \#E_i \geqslant \exp((\chi_c^- + \chi_u^- - \chi_c^\Delta - \chi_u^\Delta - r_\chi \varepsilon)(q_\chi + 1)n)$$

elements $\mathbf{y}$ in $\mathcal{Y}$. Now we arrive at the estimate

$$\begin{aligned}\mathbf{m}(\{\mathbf{t}\in R \mid F_\mathbf{s}\in\mathcal{S}_1(\chi,\varepsilon,n,p,z;S)\}) \\ &\leqslant \frac{\sum_{Y\in\mathcal{Y}} \mathbf{m}(Z(\mathbf{y}))}{\prod_{i=0}^{q_\chi} \#E_i} \\ &\leqslant \frac{\exp(((\chi_c^+ - \chi_u^- + h_\chi(\chi_c^\Delta + \chi_u^\Delta))q_\chi + (\chi_u^+ + \chi_c^{++})(q_\chi+1))n + C_\chi \varepsilon n)}{\exp((\chi_c^- + \chi_u^- - \chi_c^\Delta - \chi_u^\Delta - r_\chi\varepsilon)(q_\chi+1)n)}.\end{aligned}$$

Note that we have this estimate uniformly for the mappings F in $\mathcal{S}_1(\chi,\varepsilon,n,p,z;S)$. Put $m=q_\chi$, $T_i=\exp(\chi_c^- n)$ and $\psi_i=\psi_{n,x_i}$ for $1\leqslant i\leqslant q_\chi$ in Lemma 3.20. Then the assumption (26) holds provided that n is sufficiently large. The conclusion is that

$$\begin{aligned}\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_1(\chi,\varepsilon,n,p,z;S))\cap\mathbf{D}^{s-3}(d)) \\ &< 2^{q_\chi+1}\exp((\chi_c^{++} - \chi_c^- - \chi_u^- + (h_\chi+2)(\chi_c^\Delta + \chi_u^\Delta))q_\chi n) \\ &\quad \times \exp((\chi_c^{++} - \chi_c^- + \chi_c^\Delta + 2\chi_u^\Delta + C_\chi\varepsilon)n).\end{aligned}$$

Using the condition in the choice of q_χ, we obtain (96) for sufficiently large n, provided that we take sufficiently small ε_χ.

8. Genericity of the no flat contact condition

In this section, we consider the situation where the images of admissible curves under an iterate of a mapping $F\in\mathcal{U}$ have flat contacts with the curves in the critical set $\mathcal{C}(F)$, and investigate whether we can resolve all such flat contacts by perturbations. Our goal is the proof of Theorem 3.23, which will be carried out in the last subsection. The key idea in the proof is that the non-flatness of contacts between curves is easier to establish if we assume higher differentiability. The reader should notice that the content and the notation in this section is independent of those in the last two sections.

8.1. The jet spaces of curves

We begin with formulating a sufficient condition for the no flat contact condition in terms of *jets*. For an integer $1\leqslant q\leqslant r$ and a point $z\in M$, let Γ_z^q be the set of germs of C^q-curves $\gamma:(\mathbf{R},0)\to(M,z)$ at z. Recall that we always assume the curves to be parameterized by length. Two germs γ_1 and γ_2 in Γ^q_z are said to have contact of order q if $d(\gamma_1(t),\gamma_2(t))/|t|^q\to 0$ as $t\to 0$. This is an equivalence relation on the space Γ^q_z. The equivalence classes are called *q-jets of curve* and the quotient space is denoted by $\mathbf{J}^q\Gamma_z$. Suppose that a q-jet $\mathbf{j}$ of curve at $z\in M$ is represented by $\gamma\in\Gamma^q_z$. Then the tangent vector $d\gamma(0)/dt\in T_zM$ at z does not depend on the choice of the representative γ, and neither do the differentials $d^i\gamma(0)$, $2\leqslant i\leqslant q$, which are defined in §3.4. Thus we put

$$\mathbf{j}^{(0)}=z,\quad \mathbf{j}^{(1)}=\frac{d\gamma}{dt}(0)\quad\text{and}\quad \mathbf{j}^{(i)}=d^i\gamma(0)\quad\text{for } 2\leqslant i\leqslant q.$$

The jet space of curves of order q is the disjoint union $\mathbf{J}^q\Gamma:=\amalg_{z\in M}\mathbf{J}^q\Gamma_z$, which is equipped with the distance defined by

$$d_{\mathbf{J}}(\mathbf{j}_1,\mathbf{j}_2)=\max\{d(\mathbf{j}_1^{(0)},\mathbf{j}_1^{(0)}),\angle(\mathbf{j}_1^{(1)},\mathbf{j}_2^{(1)}),\max\{|\mathbf{j}_1^{(i)}-\mathbf{j}_2^{(i)}|\mid 2\leqslant i\leqslant q\}\}.$$

Then the mapping

$$\mathbf{j}\in\mathbf{J}^q\Gamma\longmapsto(\mathbf{j}^{(1)},(\mathbf{j}^{(i)})_{i=2}^q)\in T^1M\times\mathbf{R}^{q-1}$$

is a homeomorphism, where T^1M is the unit tangent bundle of M. Each mapping $F\in\mathcal{U}$ acts naturally on the space $\mathbf{J}^q\Gamma$. We write this action simply as

$$F:\mathbf{J}^q\Gamma\longrightarrow\mathbf{J}^q\Gamma,$$
$$[\gamma]\longmapsto[F_*\gamma].$$

For $2\leqslant q<r$, let $\mathbf{J}^q\mathcal{AC}\subset\mathbf{J}^q\Gamma$ be the compact subset of q-jets that are represented by germs of admissible curves. Lemma 3.2 tells us that $F^n(\mathbf{J}^q\mathcal{AC})\subset\mathbf{J}^q\mathcal{AC}$ for $n\geqslant n_g$.

For a C^q-curve $\gamma:I\to M$ defined on an interval I, its q-jet extension is the mapping $\mathbf{J}^q\gamma:I\to\mathbf{J}^q\Gamma$ that carries a parameter $t\in I$ to the jet in $\mathbf{J}^q\Gamma_{\gamma(t)}$ that is represented by the germ of γ at t. Recall that the critical set $\mathcal{C}(F)$ for any mapping F in $\mathcal{U}$ consists of finitely many C^{r-1}-curves. Let $\mathbf{C}(F)\subset\mathbf{J}^{r-2}\Gamma$ be the union of the images of their $(r-2)$-jet extensions:

$$\mathbf{C}(F)=\{\mathbf{J}^{r-2}\gamma(I)\mid\gamma:I\to M\text{ is a }C^{r-1}\text{-curve contained in }\mathcal{C}(F)\}.$$

LEMMA 8.1. *If a mapping $F\in\mathcal{U}$ satisfies*

$$F^n(\mathbf{J}^{r-2}\mathcal{AC})\cap\mathbf{C}(F)=\varnothing\quad\textit{for some } n\geqslant 1,\tag{97}$$

then F satisfies the no flat contact condition.

Proof. For each point in $\mathcal{C}(F)$, we can find a small C^{r-1}-coordinate neighborhood $(U,\psi:U\to\mathbf{R}^2)$ such that $\psi(\mathcal{C}(F)\cap U)$ is an interval on the x-axis $\mathbf{R}\times\{0\}$ and such that either

(a) $D\psi(\mathbf{S}^c(z))$ contains the x-axis $\mathbf{R}\times\{0\}$ for every $z\in U$, or
(b) $D\psi(\mathbf{S}^c(z))$ contains the y-axis $\{0\}\times\mathbf{R}$ for every $z\in U$.

Since the critical set $\mathcal{C}(F)$ is compact, we can cover it by finitely many coordinate neighborhoods with these properties. So, for the purpose of proving the lemma, it is enough to show the following claim for each coordinate neighborhood (U,ψ) as above: There exist $C>0$ and $n_0>0$ such that

$$\mathbf{m}_{\mathbf{R}}(\{t\in[0,a]\mid F^n(\gamma(t))\in U \text{ and } d(F^n(\gamma(t)),\mathcal{C}(F))<\varepsilon\})<C\varepsilon^{1/(r-2)}\max\{a,1\}$$

for any $a>0$, $\gamma\in\mathcal{AC}(a)$, $n\geqslant n_0$ and $\varepsilon>0$. If the condition (a) above holds, this claim is clear because the images of the admissible curves in U by the mapping ψ are curves whose slope is uniformly bounded away from 0. Thus it remains to check the claim above in the case where the condition (b) holds. To this end, it is enough to show the following lemma, because, in the case (b), the images of the admissible curves by ψ are graphs of C^{r-1}-functions whose slopes are bounded by some constant C_g.

CLAIM 8.2. *If a C^{r-1}-function φ on a compact interval $I\subset\mathbf{R}$ satisfies*

$$\max_{x\in I}\max\left\{\left|\frac{d^q\varphi}{dx^q}(x)\right| \,\middle|\, 1\leqslant q\leqslant r-1\right\}\leqslant K,$$

$$\min_{x\in I}\max\left\{\left|\frac{d^q\varphi}{dx^q}(x)\right| \,\middle|\, 1\leqslant q\leqslant r-2\right\}>\varrho$$

for some positive constants K and ϱ, then we have

$$\mathbf{m}_{\mathbf{R}}\big(\{x\in\mathbf{R}\mid |\varphi(x)|<\varepsilon\}\big)<C(r,\varrho,K,I)\,\varepsilon^{1/(r-2)} \quad \textit{for any } \varepsilon>0,$$

where $C(r,\varrho,K,I)$ is a constant that depends only on r, ϱ, K and the length of I.

We show this claim by using the following lemma [4, Lemma 5.3]:

LEMMA 8.3. *If a C^q-function h on an interval J satisfies $|d^qh(x)/dx^q|\geqslant\varrho>0$ for all $x\in J$. Then $\mathbf{m}_{\mathbf{R}}\big(\{x\in J\mid |h(x)|\leqslant\varepsilon\}\big)\leqslant 2^{q+1}(\varepsilon/\varrho)^{1/q}$ for any $\varepsilon>0$.*

Proof of Claim 8.2. Let $X\subset I$ be the set of points $x\in I$ such that $|\varphi(x)|\leqslant\frac{1}{2}\varrho$. For each point $x\in X$, there is an integer $1\leqslant m\leqslant r-2$ such that $|d^m\varphi(x)/dx^m|>\varrho$ and hence $|d^m\varphi/dx^m|\geqslant\frac{1}{2}\varrho$ on the interval $J(x):=(x-\varrho/2K, x+\varrho/2K)$. We can take points $x_i\in X$, $i=1,2,...,i_0$, so that the intervals $J(x_i)$ cover the subset X and so that the intersection multiplicity is 2, thus $i_0\leqslant 2\mathbf{m}_{\mathbf{R}}(I)/(\varrho/K)+1$. Applying Lemma 8.3 to each interval $J(x_i)$, we can see that $\mathbf{m}_{\mathbf{R}}\big(\{x\in\mathbf{R}\mid |\varphi(x)|<\varepsilon\}\big)$ is bounded by $i_0 2^{r-1}\big(\varepsilon/\frac{1}{2}\varrho\big)^{1/(r-2)}$, provided that $\varepsilon<\varrho$. This implies Claim 8.2. □

We have finished the proof of Lemma 8.1. □

8.2. Lattices in the jet space

In this subsection, we consider lattices in the space of admissible jets $\mathbf{J}^{r-2}\mathcal{AC}$ and formulate a sufficient condition for the no flat contact condition by using them. Henceforth, we fix integers $2<\nu<r\leqslant s$ satisfying the condition (3). Note that the condition (3) can be written in the form

$$(r-2)\left(r-1-\frac{r-3}{2}\right)<(r-\nu-2)\left(r-3-\frac{2s-r-\nu+1}{2\nu}\right).$$

Thus we can cover the interval $\left[\frac{1}{2}\lambda_g, 2\Lambda_g\right]$ by finitely many intervals $I(l)=(\lambda^-(l),\lambda^+(l))$, $1\leqslant l\leqslant l_0$, such that $\lambda^-(l)>\frac{1}{4}\lambda_g$ and

$$(r-2)\left(r-1-\frac{r-3}{2}\frac{\lambda^-(l)}{\lambda^+(l)}\right)<(r-\nu-2)\left(r-3-\frac{2s-r-\nu+1}{2\nu}\right).$$

For $n\geqslant 1$ and $1\leqslant l\leqslant l_0$, let $\mathbf{Q}(n,l)$ be the set of jets $\mathbf{j}$ in $\mathbf{J}^{r-2}\mathcal{AC}$ satisfying that

(Q1) the point $\mathbf{j}^{(0)}$ is contained in the lattice $\mathbf{L}(\exp(-\lambda^+(l)(r-2)n))$;

(Q2) the angle $\angle(\mathbf{j}^{(1)},\mathbf{e}^u(\mathbf{j}^{(0)}))$ is a multiple of $\exp(-\lambda^+(l)(r-3)n)$;

(Q3) $\mathbf{j}^{(q)}$ is a multiple of $\exp((-\lambda^+(l)(r-3)+\lambda^-(l)(q-1))n)$ for $2\leqslant q\leqslant r-2$.

We have

$$\#\mathbf{Q}(n,l)\leqslant C_g\exp\big((r-2)\big((r-1)\lambda^+(l)-\tfrac{1}{2}(r-3)\lambda^-(l)\big)n\big). \tag{98}$$

For integers $n\geqslant 1$, $1\leqslant l\leqslant l_0$, a mapping $F\in\mathcal{U}$ and $\sigma=0,1$, we define $V_\sigma(n,l;F)$ as the set of jets $\mathbf{j}$ in $\mathbf{J}^{r-2}\mathcal{AC}$ that satisfy

$$\exp(\lambda^-(l)n-\sigma)\leqslant|D_*F^n(\mathbf{j}^{(1)})|\leqslant\exp(\lambda^+(l)n+\sigma).$$

Then, from the choice of the numbers $\lambda^{\pm}(l)$, the subsets $V_0(n,l;F)$ for $1\leqslant l\leqslant l_0$ cover $\mathbf{J}^{r-2}\mathcal{AC}$, provided that n is larger than some constant C_g.

LEMMA 8.4. *There is a constant $B_g>1$ such that, for any jet $\mathbf{j}$ in $V_0(n,l;F)$ with $n\geqslant B_g$ and $1\leqslant l\leqslant l_0$, there exists a jet $\mathbf{i}\in\mathbf{Q}(n,l)\cap V_1(n,l;F)$ such that*

$$d_{\mathbf{J}}(F^n(\mathbf{j}),F^n(\mathbf{i}))<B_g\exp(-\lambda^+(l)(r-3)n). \tag{99}$$

Proof. Let us take a jet $\mathbf{j}\in V_0(n,l;F)$ arbitrarily. Let w be the point in the lattice $\mathbf{L}(\exp(-\lambda^+(l)(r-2)n))$ that is closest to $\mathbf{j}^{(0)}$. As $\mathbf{j}^{(1)}$ belongs to $\mathbf{S}^u(\mathbf{j}^{(0)})$, the minimum angle between $\mathbf{j}^{(1)}$ and the cone $\mathbf{S}^u(w)$ is bounded by $C_g d(\mathbf{j}^{(0)},w)$. Hence we can choose a jet $\mathbf{i}\in\mathbf{Q}(n,l)$ such that

(I1) $\mathbf{i}^{(0)}=w$ and hence $d(\mathbf{j}^{(0)},\mathbf{i}^{(0)})<\exp(-\lambda^+(l)(r-2)n)$;

(I2) $\angle(\mathbf{j}^{(1)},\mathbf{i}^{(1)})<\exp(-\lambda^+(l)(r-3)n)+C_g\exp(-\lambda^+(l)(r-2)n)$;

(I3) $|\mathbf{j}^{(q)}-\mathbf{i}^{(q)}|<\exp((-\lambda^+(l)(r-3)+\lambda^-(l)(q-1))n)$ for $2\leqslant q\leqslant r-2$.

For $0\leqslant m\leqslant n$, we put $z(m)=F^m(\mathbf{j})^{(0)}=F^m(\mathbf{j}^{(0)})$, $w(m)=F^m(\mathbf{i})^{(0)}=F^m(\mathbf{i}^{(0)})$ and

$$\Delta_m^q=\begin{cases} d(F^m(\mathbf{j})^{(0)},F^m(\mathbf{i})^{(0)})=d(z(m),w(m)) & \text{for } q=0,\\ \angle(F^m(\mathbf{j})^{(1)},F^m(\mathbf{i})^{(1)})=\angle(DF^m(\mathbf{j}^{(1)}),DF^m(\mathbf{i}^{(1)})) & \text{for } q=1,\\ |F^m(\mathbf{j})^{(q)}-F^m(\mathbf{i})^{(q)}| & \text{for } 2\leqslant q\leqslant r-2.\end{cases}$$

In order to prove the inequality (99), it is enough to show that

$$\Delta_n^q\leqslant C_g\exp(-\lambda^+(l)(r-3)n)\quad \text{for } 0\leqslant q\leqslant r-2.$$

First we prove that

$$\Delta_m^0\leqslant 2\|DF_{z(0)}^m\|\,\Delta_0^0\leqslant C_g\exp(-\lambda^+(l)(r-3)n) \tag{100}$$

for $1\leqslant m<n$. As $\mathbf{j}\in V_0(n,l;F)$, we have

$$\begin{aligned}\|DF_{z(k)}^{m-k}\|&\leqslant C_g D_*F^{m-k}(DF^k(\mathbf{j}^{(1)}))\\ &\leqslant\frac{C_g D_*F^n(\mathbf{j}^{(1)})}{D_*F^{n-m}(DF^m(\mathbf{j}^{(1)}))\,D_*F^k(\mathbf{j}^{(1)})}\\ &\leqslant C_g\exp(\lambda^+(l)n-\lambda_g(n-m+k))\end{aligned} \tag{101}$$

for $0\leqslant k\leqslant m\leqslant n$. So the second inequality in (100) follows from the condition (I1). We prove the first inequality in (100) by induction on $1\leqslant m<n$. Using the simple estimate

$$\|\exp_{z(m)}^{-1}(w(m))-DF_{z(m-1)}(\exp_{z(m-1)}^{-1}(w(m-1)))\|\leqslant C_g(\Delta_{m-1}^0)^2$$

repeatedly, we can get the following inequality for $\Delta_m^0=\|\exp_{z(m)}^{-1}(w(m))\|$:

$$\Delta_m^0\leqslant\|DF_{z(0)}^m\|\,\Delta_0^0+C_g\sum_{k=0}^{m-1}\|DF_{z(k+1)}^{m-k-1}\|\,(\Delta_k^0)^2\quad\text{for } 0\leqslant m\leqslant n. \tag{102}$$

Note that we have, from (6),

$$\begin{aligned}\|DF_{z(k+1)}^{m-k-1}\|\cdot\|DF_{z(0)}^k\|&\leqslant C_g D_*F^{m-k-1}(DF^{k+1}(\mathbf{e}^u(z_0)))\,D_*F^k(\mathbf{e}^u(z_0))\\ &\leqslant C_g\frac{D_*F^m(\mathbf{e}^u(z_0))}{D_*F(DF^k(\mathbf{e}^u(z_0)))}\leqslant C_g\|DF_{z(0)}^m\|\end{aligned} \tag{103}$$

for $0\leqslant k\leqslant m-1$. Consider an integer $0\leqslant m_0\leqslant n$ and suppose that the left inequality in (100) holds for $0\leqslant m<m_0$. Then, using the estimates (101) and (103) in (102), we obtain

$$\begin{aligned}\Delta_{m_0}^0&\leqslant\|DF_{z(0)}^{m_0}\|\,\Delta_0^0+C_g\sum_{k=0}^{m_0-1}\|DF_{z(k+1)}^{m_0-k-1}\|\cdot 2\|DF_{z(0)}^k\|\,\Delta_0^0\Delta_k^0\\ &\leqslant\|DF_{z(0)}^{m_0}\|\,\Delta_0^0(1+C_g n\exp(-\lambda^+(l)(r-3)n)).\end{aligned}$$

This implies the first inequality in (100) for $m=m_0$, provided that n is larger than some constant C_g. Thus we can obtain (100) for $1\leqslant m\leqslant n$ by induction.

Next we estimate Δ^1_m for $0\leqslant m\leqslant n$. We have

$$\Delta^1_m \leqslant \angle(DF^m_{z(0)}(\mathbf{j}^{(1)}), DF^m_{z(0)}(\mathbf{i}^{(1)})) + \sum_{k=0}^{m-1} \angle(DF^{m-k}_{z(k)}(DF^k_{w(0)}(\mathbf{i}^{(1)})), DF^{m-k-1}_{z(k+1)}(DF^{k+1}_{w(0)}(\mathbf{i}^{(1)}))),$$

where we omit the operations of parallel translation (see the remark given in the proof of Lemma 5.1). For $0\leqslant k<n$, we have $DF^k_{w(0)}(\mathbf{i}^{(1)})\in \mathbf{S}^u(w(k))$ and $d(z(k),w(k))=\Delta^0_k\leqslant C_g\exp(-\lambda^+(l)(r-3)n)$. Hence the parallel translation of $DF^k_{w(0)}(\mathbf{i}^{(1)})$ to $z(k)$ does not belong to the central cone $\mathbf{S}^c(z(k))$, provided that n is larger than some constant C_g. Using (4), we can obtain

$$\begin{aligned}
&\angle(DF^{m-k}_{z(k)}(DF^k_{w(0)}(\mathbf{i}^{(1)})), DF^{m-k-1}_{z(k+1)}(DF^{k+1}_{w(0)}(\mathbf{i}^{(1)})))\\
&\leqslant A_g \frac{|D^*F^{m-k-1}(DF^{k+1}(\mathbf{j}^{(1)}))|}{D_*F^{m-k-1}(DF^{k+1}(\mathbf{j}^{(1)}))} \angle(DF_{z(k)}(DF^k_{w(0)}(\mathbf{i}^{(1)})), DF_{w(k)}(DF^k_{w(0)}(\mathbf{i}^{(1)})))\\
&\leqslant C_g\exp(-\lambda_g(m-k-1))\Delta^0_k\\
&< C_g\exp(-\lambda_g(m-k-1)-\lambda^+(l)(r-3)n).
\end{aligned}$$

Likewise we can obtain $\angle(DF^m_{z(0)}(\mathbf{j}^{(1)}), DF^m_{z(0)}(\mathbf{i}^{(1)}))\leqslant C_g\exp(-\lambda_g m)\Delta^1_0$. Therefore, by condition (I2), we conclude that

$$\begin{aligned}
\Delta^1_m &\leqslant C_g\exp(-\lambda_g m)\Delta^1_0+\sum_{k=0}^{m-1} C_g\exp(-\lambda_g(m-k-1)-\lambda^+(l)(r-3)n)\\
&\leqslant C_g\exp(-\lambda^+(l)(r-3)n).
\end{aligned}$$

Finally, we estimate Δ^q_n for $2\leqslant q\leqslant r$. From the formula (10), we can see that

$$\Delta^q_m \leqslant \frac{|D^*F(DF^{m-1}(\mathbf{j}^{(1)}))|}{D_*F(DF^{m-1}(\mathbf{j}^{(1)}))^q}\Delta^q_{m-1}+C_g\sum_{0\leqslant d<q}\Delta^d_{m-1}. \tag{104}$$

Consider this inequality for $m=n$ and estimate the right-hand side by using (104) recursively as long as there exist terms Δ^q_m with $q>1$ or $m>0$ on the right-hand side. Then we see that Δ^q_n is bounded by

$$\begin{aligned}
&\frac{|D^*F^n(\mathbf{j}^{(1)})|}{D_*F^n(\mathbf{j}^{(1)})^q}\Delta^q_0 + C_g\sum_{1<d<q}\ \sum_{\substack{0=n_d\leqslant n_{d+1}\leqslant\ldots\\ \leqslant n_q<n_{q+1}=n+1}}\ \prod_{l=d}^{q}\ \prod_{n_l\leqslant j<n_{l+1}-1}\frac{|D^*F(F^j(\mathbf{j})^{(1)})|}{D_*F(F^j(\mathbf{j})^{(1)})^l}\Delta^d_0\\
&\quad + C_g\sum_{\substack{d=0,1\\ 0\leqslant m<n}}\ \sum_{\substack{m=n_d\leqslant n_{d+1}\leqslant\ldots\\ \leqslant n_q<n_{q+1}=n+1}}\ \prod_{l=d}^{q}\ \prod_{n_l\leqslant j<n_{l+1}-1}\frac{|D^*F(F^j(\mathbf{j})^{(1)})|}{D_*F(F^j(\mathbf{j})^{(1)})^l}\Delta^d_m.
\end{aligned} \tag{105}$$

Note that, for any sequence $m=n_d \leqslant n_{d+1} \leqslant \ldots \leqslant n_q < n_{q+1} = n+1$ with $q \leqslant r$, we have

$$\prod_{l=d}^{q} \prod_{n_l \leqslant j < n_{l+1}-1} \frac{|D^*F(F^j(\mathbf{j})^{(1)})|}{D_*F(F^j(\mathbf{j})^{(1)})^l} \leqslant \frac{\exp(-\lambda_g(n-m-q)+c_g q)}{D_*F^{n-m}(F^m(\mathbf{j})^{(1)})^{d-1}\exp(-q^2\Lambda_g)}$$
$$\leqslant C_g \frac{\exp(-\lambda_g(n-m))}{D_*F^{n-m}(F^m(\mathbf{j})^{(1)})^{d-1}}.$$

Hence it follows from (105) that

$$\begin{aligned}\Delta_n^q &\leqslant \frac{\exp(-\lambda_g n)}{D_*F^n(\mathbf{j}^{(1)})^{q-1}}\Delta_0^q + C_g n^q \sum_{1<d<q} \frac{\exp(-\lambda_g n)}{D_*F^n(\mathbf{j}^{(1)})^{d-1}}\Delta_0^d \\ &\quad + C_g \sum_{0\leqslant m<n} (n-m)^q \exp(-\lambda_g(n-m))(\Delta_m^0+\Delta_m^1) \\ &\leqslant C_g \max_{0\leqslant m<n}(\Delta_m^0+\Delta_m^1) + C_g \sum_{1<d\leqslant q} \exp(-(d-1)\lambda^-(l)n)\Delta_0^d,\end{aligned}$$

where the second inequality follows from the fact that the jet $\mathbf{j}$ belongs to $V_0(n,l;F)$. Using the estimates on Δ_m^0 and Δ_m^1, and the condition (I3) in the inequality above, we can conclude that

$$\Delta_n^q < C_g \exp(-\lambda^+(l)(r-3)n) \quad \text{for } 2 \leqslant q \leqslant r-2.$$

We have proved the inequality (99). The jet $\mathbf{i}$ belongs to $V_1(n,l;F)$ because

$$\log \frac{D_*F^n(\mathbf{e}^u(\mathbf{i}^{(0)}))}{D_*F^n(\mathbf{e}^u(\mathbf{j}^{(0)}))} \leqslant C_g \sum_{m=0}^{n-1}(\Delta_m^0+\Delta_m^1) \leqslant C_g n \exp(-\lambda^+(l)(r-3)n) < 1,$$

provided that n is larger than some constant C_g. □

For integers $n \geqslant 1$, $1 \leqslant l \leqslant l_0$ and a jet $\mathbf{j} \in \mathbf{Q}(n,l)$, let $\mathcal{S}_2(n,l,\mathbf{j})$ be the set of mappings $F \in \mathcal{U}$ such that $\mathbf{j} \in V_1(n,l;F)$ and

$$d_{\mathbf{J}}(F^n(\mathbf{j}), \mathbf{C}(F)) \leqslant 2B_g \exp(-\lambda^+(l)n(r-3)).$$

Then the last lemma implies the following result:

Corollary 8.5. *If there exists $n \geqslant B_g$ such that $F \notin \mathcal{S}_2(n,l,\mathbf{j})$ for all $1 \leqslant l \leqslant l_0$ and $\mathbf{j} \in \mathbf{Q}(n,l)$, then F satisfies the no flat contact condition.*

In the remaining part of this section, we shall estimate the measure of the subsets $\mathcal{S}_2(n,l,\mathbf{j})$ for $\mathbf{j} \in \mathbf{Q}(n,l)$ by using Lemma 3.20.

8.3. Perturbations

In this subsection, we introduce some families of mappings and give a few estimates on the variation of the images of jets under the iterates of mappings in the families. In the argument below, we fix $1\leqslant l\leqslant l_0$ and put

$$\delta_n=\exp\left(-\frac{\lambda^+(l)n}{\nu}\right)\quad\text{for } n\geqslant 1.$$

For $1\leqslant q\leqslant r-2$, we fix a C^∞-function $\psi_q\colon\mathbf{R}^2\to\mathbf{R}$ such that

$$\psi_q(x,y)=\begin{cases}x^q/q! & \text{for } (x,y)\in\mathbf{B}(0,\frac{1}{10}),\\ 0 & \text{for } (x,y)\notin\mathbf{B}(0,1).\end{cases}$$

Remark. We can take the functions ψ_q so that their C^r-norm is bounded by some constant C_g.

For each point $\zeta\in M$, we consider an isometric embedding

$$\varphi_\zeta\colon\{w\in\mathbf{R}^2\mid \|w\|<\tfrac{1}{5}\}\longrightarrow\mathbf{T}$$

that carries the origin to the point ζ and the x-axis $\mathbf{R}\times\{0\}$ to $\mathbf{E}^u(\zeta)$.

Recall that we took the subset $\mathcal{U}$ of mappings as a neighborhood of a C^r-mapping $F_\sharp$ in §3.2. For positive integers n, $1\leqslant q\leqslant r-2$ and a point ζ in M, we define a C^∞-mapping $\psi_{q,n,\zeta}\colon M\to\mathbf{R}^2$ by

$$\psi_{q,n,\zeta}(z)=\begin{cases}\delta_n^s\psi_q(\varphi_\zeta^{-1}(z)/\delta_n)\,\mathbf{e}^c(F_\sharp(\zeta)), & \text{if } d(z,\zeta)<\delta_n,\\ 0, & \text{otherwise,}\end{cases}$$

where $\mathbf{e}^c(\cdot)$ is either of the two unit tangent vectors in the central subspace $\mathbf{E}^c(\cdot)$. Note that, for any mapping $F\in\mathcal{U}$, the parallel translation of the vector $\mathbf{e}^c(F_\sharp(z))$ to $F(z)$ is contained in $\mathbf{S}^c(F(z))$ from the choice of $\mathcal{U}$ in §3.2.

For a positive integer n, a mapping $F\in\mathcal{U}$ and a point ζ in M, we define

$$F_{\mathbf{t}}(z)=F(z)+\sum_{q=\nu+1}^{r-2}t_q\psi_{q,n,\zeta}(z)\colon M\longrightarrow\mathbf{T},\tag{106}$$

where $\mathbf{t}=(t_{\nu+1},t_{\nu+2},\dots,t_{r-2})$ is the parameter that ranges over $R=[-1,1]^{r-2-\nu}$. This is the family of mappings that we are going to consider.

Remark. The purpose of considering the family above is to move the images $F_{\mathbf{t}}^n(\mathbf{j})$ of the jets $\mathbf{j}\in\mathbf{Q}(n,l)$ by choosing the point ζ appropriately. As it will turn out, we can keep control of the coordinates $F_{\mathbf{t}}^n(\mathbf{j})^{(q)}$ with $q\geqslant\nu+1$, but *not* of those with $0\leqslant q\leqslant\nu$.

This is the reason why we restricted the range of q between $\nu+1$ and $r-2$ in (106). Note that, if we take smaller ν, we can keep control of more coordinates but the magnitude of the perturbation becomes smaller. Thus, we have to choose a good value for ν. The inequality (3) is related to this choice.

Obviously we have

$$d_{C^q}(F_{\mathbf{t}}, F) \leqslant C_g \delta_n^{s-q} \quad \text{and} \quad \|\partial_{\mathbf{t}} F_{\mathbf{t}}\|_{C^q} \leqslant C_g \delta_n^{s-q} \tag{107}$$

for $0 \leqslant q \leqslant r$ and $\mathbf{t} \in R$. In particular, $F_{\mathbf{t}}(M) \subset M$ if n is sufficiently large.

We consider a jet $\mathbf{j} \in \mathbf{Q}(n,l) \cap V_1(n,l;F)$ and give some estimates on the variation of the image $F_{\mathbf{t}}^n(\mathbf{j})$. We begin with the estimate on the position $F_{\mathbf{t}}^n(\mathbf{j})^{(0)}$.

LEMMA 8.6. *We have, for* $0 \leqslant m \leqslant n$ *and* $\mathbf{t} \in R$,

$$d(F_{\mathbf{t}}^m(\mathbf{j}^{(0)}), F^m(\mathbf{j}^{(0)})) < C_g \|DF_{\mathbf{j}^{(0)}}^m\| \delta_n^s \leqslant C_g \delta_n^{s-\nu}$$

and

$$\|\partial_{\mathbf{t}} F_{\mathbf{t}}^m(\mathbf{j}^{(0)})\| < C_g \|DF_{\mathbf{j}^{(0)}}^m\| \delta_n^s \leqslant C_g \delta_n^{s-\nu},$$

provided that n *is larger than some constant* C_g.

Proof. The following argument is a modification of that in the former part of the proof of Lemma 8.4. We put $z(m) = F^m(\mathbf{j}^{(0)})$, $w(m) = F_{\mathbf{t}}^m(\mathbf{j}^{(0)})$ and $\Delta_m = d(z(m), w(m))$ for $0 \leqslant m \leqslant n$, so that $\Delta_0 = 0$. Using the simple estimate

$$\|\exp_{z(m)}^{-1}(w(m)) - (DF)_{z(m-1)}(\exp_{z(m-1)}^{-1}(w(m-1)))\| \leqslant C_g(\delta_n^s + (\Delta_{m-1})^2)$$

repeatedly, we obtain

$$\Delta_m \leqslant \sum_{k=0}^{m-1} \|(DF^{m-k-1})_{z(k+1)}\| C_g(\delta_n^s + (\Delta_k)^2) \tag{108}$$

for $0 \leqslant m \leqslant n$. Consider an integer $0 \leqslant m_0 \leqslant n$ and a positive number K, and suppose that we have

$$\Delta_m < K \|(DF^m)_{z(0)}\| \delta_n^s \tag{109}$$

for $0 \leqslant m < m_0$. Then, using this, the inequality (103) and the simple estimate

$$C_g^{-1} \exp(\lambda_g k) \leqslant \|DF_{z(0)}^k\| \leqslant C_g \|DF_{z(0)}^n\| \leqslant C_g \delta_n^{-\nu} \quad \text{for } 0 \leqslant k \leqslant m \leqslant n$$

on the right-hand side of the inequality (108) for $m=m_0$, we obtain

$$\Delta_{m_0} \leqslant C_g \|(DF^{m_0})_{z(0)}\| \sum_{k=0}^{m-1} (\delta_n^s \|DF^k_{z(0)}\|^{-1} + K^2 \delta_n^{2s} \|DF^k_{z(0)}\|)$$
$$\leqslant C_g \|(DF^{m_0})_{z(0)}\| \, \delta_n^s \sum_{k=0}^{m-1} (\exp(-\lambda_g k) + K^2 \delta_n^{s-\nu}).$$

This implies (109) for $m=m_0$, provided that K and n are larger than some constant C_g. Thus we can obtain the first claim of the lemma by induction on m.

Put $\Delta'_m = \partial_{\mathbf{t}} F^m_{\mathbf{t}}(\mathbf{j}^{(0)})$ for $0 \leqslant m \leqslant n$. Using the simple estimate

$$\|\Delta'_m - (DF)_{z(m-1)} \Delta'_{m-1}\| \leqslant C_g (\delta_n^s + \Delta_{m-1} \|\Delta'_{m-1}\|)$$

repeatedly, we obtain

$$\Delta'_m \leqslant \sum_{k=0}^{m-1} \|(DF^{m-k-1})_{z(k+1)}\| \, C_g (\delta_n^s + \Delta_k \|\Delta'_k\|).$$

From this and the estimates on Δ_m that we have proved above, we can obtain the second claim of the lemma by induction on m, in a similar manner as above. □

Next we give the estimates on $\partial_{\mathbf{t}} F^n_{\mathbf{t}}(\mathbf{j})^{(q)}$ for $1 \leqslant q \leqslant r-2$. We use the notation ∂_p for the differentiation by the parameter t_p. For integers p and q satisfying $\nu+1 \leqslant p \leqslant r-2$ and $1 \leqslant q \leqslant r-2$, and for a jet $\mathbf{i} \in \mathbf{J}^{r-2}\mathcal{A}C$ and $\mathbf{t} \in R$, we define

$$\beta_p^{(q)}(\mathbf{i}, \mathbf{t}) = \pm \frac{\sin(\angle(\mathbf{e}^c(F_\sharp(z)), F_{\mathbf{t}}(\mathbf{i}^{(1)}))) \, \|\partial_p((D^q F_{\mathbf{t}})_{\mathbf{i}^{(0)}}(\mathbf{i}^{(1)}, \mathbf{i}^{(1)}, \dots, \mathbf{i}^{(1)}))\|}{D_* F_{\mathbf{t}}(\mathbf{i}^{(1)})^q},$$

where $(D^q F_{\mathbf{t}})_z \colon \bigotimes^q T_z M \to T_{F(z)} M$ is the qth differential of $F_{\mathbf{t}}$ at z, and the sign on the right-hand side will be chosen appropriately in the argument below. We have

$$|\beta_p^{(q)}(\mathbf{i}, \mathbf{t})| \leqslant C_g \delta_n^{s-q}. \tag{110}$$

LEMMA 8.7. *There exists a positive constant C_g such that, if $n \geqslant C_g$, we have*

$$\left| \partial_p (F^m_{\mathbf{t}}(\mathbf{j})^{(q)}) - \sum_{k=0}^{m-1} \frac{D^* F^{m-k-1}_{\mathbf{t}}(F^{k+1}_{\mathbf{t}}(\mathbf{j})^{(1)})}{D_* F^{m-k-1}_{\mathbf{t}}(F^{k+1}_{\mathbf{t}}(\mathbf{j})^{(1)})^q} \beta_p^{(q)}(F^k_{\mathbf{t}}(\mathbf{j}), \mathbf{t}) \right| < C_g \delta_n^{s-q+1}$$

for any $\nu+1 \leqslant q \leqslant r-2$, $\nu+1 \leqslant p \leqslant r-2$, $\mathbf{t} \in R$ and $0 \leqslant m \leqslant n$, provided that we choose the sign in the definition of $\beta_p^{(q)}(F^k_{\mathbf{t}}(\mathbf{j}), \mathbf{t})$ appropriately.

Proof. Fix $\nu+1 \leqslant p \leqslant r-2$ arbitrarily. For $0 \leqslant q \leqslant r-2$ and $0 \leqslant m \leqslant n$, we put

$$\Delta_m^{(q)} = \begin{cases} \|\partial_p F^m_{\mathbf{t}}(\mathbf{j})^{(0)}\|, & \text{if } q=0, \\ \partial_p \angle(F^m_{\mathbf{t}}(\mathbf{j})^{(1)}, v_0), & \text{if } q=1, \\ \partial_p (F^m_{\mathbf{t}}(\mathbf{j})^{(q)}), & \text{if } q \geqslant 2, \end{cases}$$

where v_0 is some fixed vector. For $0<m\leqslant n$, we have

$$\left|\Delta_m^{(1)}-\frac{D^*F_{\mathbf{t}}(F_{\mathbf{t}}^{m-1}(\mathbf{j})^{(1)})}{D_*F_{\mathbf{t}}(F_{\mathbf{t}}^{m-1}(\mathbf{j})^{(1)})}\Delta_{m-1}^{(1)}-\beta_p^{(1)}(F_{\mathbf{t}}^{m-1}(\mathbf{j}),\mathbf{t})\right|\leqslant C_g\Delta_{m-1}^{(0)}\leqslant C_g\delta_n^{s-\nu},$$

where the second inequality follows from Lemma 8.6. From this inequality and the estimate (110) for $q=1$, we see that

$$|\Delta_m^{(1)}|\leqslant C_g\sum_{k=0}^{m-1}\frac{|D^*F_{\mathbf{t}}^{m-k}(F_{\mathbf{t}}^k(\mathbf{j})^{(1)})|}{D_*F_{\mathbf{t}}^{m-k}(F_{\mathbf{t}}^k(\mathbf{j})^{(1)})}(\beta_p^{(1)}(F_{\mathbf{t}}^k(\mathbf{j}),\mathbf{t})+\delta_n^{s-\nu})<C_g\delta_n^{s-\nu}$$

for $0\leqslant m\leqslant n$. Recall the formula (10) and the remark after it. By differentiating both sides of (10) with F replaced by $F_{\mathbf{t}}$ and using (107), we obtain

$$\left|\Delta_m^{(q)}-\frac{D^*F_{\mathbf{t}}(F_{\mathbf{t}}^{m-1}(\mathbf{j})^{(1)})}{D_*F_{\mathbf{t}}(F_{\mathbf{t}}^{m-1}(\mathbf{j})^{(1)})^q}\Delta_{m-1}^{(q)}-\beta_p^{(q)}(F_{\mathbf{t}}^m(\mathbf{j}),\mathbf{t}))\right|\leqslant C_g\delta_n^{s-q+1}+C_g\sum_{0\leqslant d<q}\Delta_{m-1}^{(d)}\qquad(111)$$

for $2\leqslant q\leqslant r-2$ and $0\leqslant m\leqslant n$. In particular, we have, from (110),

$$\left|\Delta_m^{(q)}-\frac{D^*F_{\mathbf{t}}(F_{\mathbf{t}}^{m-1}(\mathbf{j})^{(1)})}{D_*F_{\mathbf{t}}(F_{\mathbf{t}}^{m-1}(\mathbf{j})^{(1)})^q}\Delta_{m-1}^{(q)}\right|\leqslant C_g\delta_n^{s-q+1}+C_g\sum_{0\leqslant d<q}\Delta_{m-1}^{(d)}$$

for $2\leqslant q\leqslant r-2$ and $0\leqslant m\leqslant n$. Using this inequality repeatedly, we reach

$$\begin{aligned}|\Delta_m^{(q)}|&\leqslant C_g\sum_{k=1}^{m}\frac{|D^*F_{\mathbf{t}}^{m-k}(F_{\mathbf{t}}^k(\mathbf{j})^{(1)})|}{D_*F_{\mathbf{t}}^{m-k}(F_{\mathbf{t}}^k(\mathbf{j})^{(1)})^q}\left(\delta_n^{s-q}+\sum_{0\leqslant d<q}\Delta_{k-1}^{(d)}\right)\\&\leqslant C_g\left(\delta_n^{s-q}+\max_{0\leqslant d<q}\max_{0<k<m}\Delta_k^{(d)}\right).\end{aligned}$$

Hence, we can show the estimate $|\Delta_m^{(q)}|\leqslant C_g\delta_n^{s-\nu}$ for $2\leqslant q\leqslant\nu$ and $0\leqslant m\leqslant n$ by induction on q, by Lemma 8.6 and the estimate on $|\Delta_m^{(1)}|$ above.

Next, using the inequality (111) repeatedly, we can see that the left-hand side of the inequality in the lemma is bounded by

$$C_g\sum_{k=1}^{m}\frac{|D^*F_{\mathbf{t}}^{m-k}(F_{\mathbf{t}}^k(\mathbf{j})^{(1)})|}{D_*F_{\mathbf{t}}^{m-k}(F_{\mathbf{t}}^k(\mathbf{j})^{(1)})^q}\left(\delta_n^{s-q+1}+\sum_{0\leqslant d<q}\Delta_{k-1}^{(d)}\right).$$

By induction on $\nu+1\leqslant q\leqslant r-2$, we obtain the inequality in the lemma. □

Note that, for any C^r-mapping $G: M\to M$ such that $d_{C^r}(G,F_\sharp)<2\varrho_g$, the level curves of the function $\det G: z\mapsto\det DG_z$ are regular in the neighborhood $\mathbf{B}(\mathcal{C}(G),\varrho_g)$ of the critical set $\mathcal{C}(G)$, from the choice of the constant ϱ_g in §3.2. For a point $w\in\mathbf{B}(\mathcal{C}(G),\varrho_g)$, let $\mathbf{c}(w;G)$ be the $(r-2)$-jet at w that is given by the level curve passing through w.

Suppose that a jet $\mathbf{j}\in\mathbf{J}^{r-2}\mathcal{AC}$ satisfies, for all $\mathbf{t}\in R$,

(V1) $d(F_{\mathbf{t}}^{n-1}(\mathbf{j})^{(0)},\zeta)<\frac{1}{10}\delta_n$;

(V2) $d(F_{\mathbf{t}}^{n-1}(\mathbf{j})^{(0)},\mathcal{C}(F_{\mathbf{t}}))>3\delta_n$;

(V3) $d(F_{\mathbf{t}}^{n}(\mathbf{j})^{(0)},\mathcal{C}(F_{\mathbf{t}}))<\delta_n$.

From the condition (V3), we can define the mapping $\Psi: R \to \mathbf{R}^{r-\nu-2}$ by

$$\Psi(\mathbf{t}) = \left\{ \frac{F_{\mathbf{t}}^n(\mathbf{j})^{(q)} - \mathbf{c}(F_{\mathbf{t}}^n(\mathbf{j})^{(0)}; F_{\mathbf{t}})^{(q)}}{\delta_n^{s-q}} \right\}_{q=\nu+1}^{r-2},$$

provided that n is so large that $\delta_n < \varrho_g$. The next lemma is the goal of this subsection:

LEMMA 8.8. *If the conditions* (V1), (V2) *and* (V3) *hold for all* $\mathbf{t} \in R$, *then the mapping* Ψ *is a diffeomorphism and* $|\det D\Psi(\mathbf{t})|$ *is bounded from below by a constant* C_g^{-1}, *provided that* n *is larger than some constant* C_g.

Proof. From the condition (V1) and the definition of the family $F_{\mathbf{t}}$, we have

$$\beta_p^{(q)}(F_{\mathbf{t}}^{n-1}(\mathbf{j}), \mathbf{t}) = 0 \qquad \text{for } q > p,$$
$$|\beta_p^{(q)}(F_{\mathbf{t}}^{n-1}(\mathbf{j}), \mathbf{t})| \geqslant C_g^{-1}\delta_n^{s-q} \quad \text{for } q = p,$$

in addition to (110). We show that

$$\left| \sum_{k=0}^{n-2} \frac{D^* F_{\mathbf{t}}^{n-k-1}(F_{\mathbf{t}}^{k+1}(\mathbf{j})^{(1)})}{D_* F_{\mathbf{t}}^{n-k-1}(F_{\mathbf{t}}^{k+1}(\mathbf{j})^{(1)})^q} \beta_p^{(q)}(F_{\mathbf{t}}^k(\mathbf{j}), \mathbf{t}) \right| < C_g \delta_n^{s-q+1}. \tag{112}$$

Suppose that $\beta_p^{(q)}(F_{\mathbf{t}}^k(\mathbf{j}), \mathbf{t}) \neq 0$ for some integer $0 \leqslant k \leqslant m-2$. Then we find that $d(F_{\mathbf{t}}^k(\mathbf{j})^{(0)}, \zeta) < \delta_n$ and

$$\begin{aligned} d(F_{\mathbf{t}}^{k+1}(\mathbf{j})^{(0)}, \mathcal{C}(F_{\mathbf{t}})) &\leqslant d(F_{\mathbf{t}}^{k+1}(\mathbf{j})^{(0)}, F_{\mathbf{t}}(\zeta)) + d(F_{\mathbf{t}}(\zeta), F_{\mathbf{t}}^n(\mathbf{j})^{(0)}) + d(F_{\mathbf{t}}^n(\mathbf{j})^{(0)}, \mathcal{C}(F_{\mathbf{t}})) \\ &< C_g \delta_n \end{aligned}$$

from (V1) and (V3). This and (5) imply that $|D^* F_{\mathbf{t}}(F_{\mathbf{t}}^{k+1}(\mathbf{j})^{(1)})| < C_g \delta_n$, and hence

$$\left| \frac{D^* F_{\mathbf{t}}^{m-k-1}(F_{\mathbf{t}}^{k+1}(\mathbf{j})^{(1)})}{D_* F_{\mathbf{t}}^{m-k-1}(F_{\mathbf{t}}^{k+1}(\mathbf{j})^{(1)})^q} \beta_p^{(q)}(F_{\mathbf{t}}^k(\mathbf{j}), \mathbf{t}) \right| \leqslant C_g \delta_n^{s-q+1} \exp(-\lambda_g(m-k-1) + 2c_g).$$

Therefore we obtain (112).

The jet $\mathbf{c}(w; F_{\mathbf{t}})$ for $w \in \mathbf{B}(\mathcal{C}(F), \delta_n)$ does not depend on the parameter $\mathbf{t} \in R$ because $\mathbf{B}(\zeta, \delta_n) \cap \mathbf{B}(\mathcal{C}(F), \delta_n) = \varnothing$ from (V1) and (V2). So we have

$$\|\partial_p(\mathbf{c}(F_{\mathbf{t}}^n(\mathbf{j})^{(0)}; F_{\mathbf{t}})^{(q)})\| < C_g \|\partial_p(F_{\mathbf{t}}^n(\mathbf{j})^{(0)})\| < C_g \delta_n^{s-\nu} \quad \text{for } \nu+1 \leqslant p, q \leqslant r-2$$

by Lemma 8.6. From (112) and Lemma 8.7, it follows that

$$|\partial_p(F_{\mathbf{t}}^m(\mathbf{j})^{(q)} - \mathbf{c}(F_{\mathbf{t}}^n(\mathbf{j})^{(0)}; F_{\mathbf{t}})^{(q)}) - \beta_p^{(q)}(F_{\mathbf{t}}^{m-1}(\mathbf{j}), \mathbf{t})| < C_g \delta_n^{s-q+1}.$$

Let $D\Psi(\mathbf{t})_{q,p}$ be the (q,p)-entry of the representation matrix of $D\Psi(\mathbf{t})$ with respect to the standard basis of $\mathbf{R}^{r-2-\nu}$. Then, from the estimates above, we have

$$|D\Psi(\mathbf{t})_{q,p}| < C_g \delta_n \quad \text{if } q > p,$$
$$|D\Psi(\mathbf{t})_{q,p}| < C_g \qquad \text{if } q \leqslant p,$$
$$|D\Psi(\mathbf{t})_{q,p}| > C_g^{-1} \quad \text{if } q = p.$$

Now we can conclude the lemma by an elementary argument. □

8.4. Resolution of the flat contacts

In this subsection, we prove Theorem 3.23. Until the last part of the proof, we fix $1 \leqslant l \leqslant l_0$ and put $\delta_n = \exp(-\lambda^+(l)n/\nu)$ for $n \geqslant 1$ as in the last subsection. Let n be a large integer, ζ a point in the lattice $\mathbf{L}(\frac{1}{20}\delta_n)$ and $\mathbf{j}$ a jet in $\mathbf{Q}(n,l)$. Let $Y_0(n,l,\mathbf{j},\zeta)$ (resp. $Y_1(n,l,\mathbf{j},\zeta)$), be the set of mappings $F \in C^r(M,M)$ that satisfy

$$F^{n-1}(\mathbf{j})^{(0)} \in \mathbf{B}(\zeta, \tfrac{1}{20}\delta_n) \quad \big(\text{resp. } F^{n-1}(\mathbf{j})^{(0)} \in \mathbf{B}(\zeta, \tfrac{1}{5}\delta_n)\big). \tag{113}$$

Below we estimate

$$\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_2(n,l,\mathbf{j}) \cap Y_0(n,l,\mathbf{j},\zeta)) \cap \mathbf{D}^r(d)) \quad \text{for } G \in C^r(M,\mathbf{T}) \text{ and } d > 0,$$

where Φ_G and $\mathbf{D}^r(d)$ are defined by (2) and (25), respectively.

Take a mapping F in $\mathcal{S}_2(n,l,\mathbf{j}) \cap Y_0(n,l,\mathbf{j},\zeta)$ arbitrarily and consider the family $F_{\mathbf{t}}$ defined by (106) in the last subsection. Note that the jet $\mathbf{j}$ belongs to $V_1(n,l;F)$ from the definition of $\mathcal{S}_2(n,l,\mathbf{j})$. We check that the conditions (V1), (V2) and (V3) hold for $\mathbf{t} \in R$, provided that n is larger than some constant C_g. Since F belongs to $\mathcal{S}_2(n,l,\mathbf{j})$, there exists a point $w_0 \in \mathcal{C}(F)$ such that

$$d_{\mathbf{J}}(F^n(\mathbf{j}), \mathbf{c}(w_0;F)) \leqslant 2B_g\delta_n^{(r-3)\nu}. \tag{114}$$

In particular, we have $d(F^n(\mathbf{j})^{(0)}, w_0) < \varrho_g$ and $\angle(F^n(\mathbf{j})^{(1)}, \mathbf{c}(w_0;F)^{(1)}) < \varrho_g$, provided that n is larger than some constant C_g. It follows from the condition (C5) in the choice of the constant ϱ_g in §3.2 that

$$d(F^{n-1}(\mathbf{j})^{(0)}, \mathcal{C}(F)) > \varrho_g. \tag{115}$$

Using (113), we can see that

$$d(\zeta, \mathcal{C}(F)) \geqslant d(F^{n-1}(\mathbf{j})^{(0)}, \mathcal{C}(F)) - d(F^{n-1}(\mathbf{j})^{(0)}, \zeta) > \varrho_g - 2B_g\delta_n^{(r-3)\nu} > 4\delta_n,$$

provided that n is larger than some constant C_g. This implies that the critical set $\mathcal{C}(F_{\mathbf{t}})$ does not depend on $\mathbf{t} \in R$. Hence (V1), (V2) and (V3) follow from (113), (114), (115) and Lemma 8.6, provided that n is larger than some constant C_g.

Let $\Psi: R \to \mathbf{R}^{r-\nu-2}$ be the mapping defined in the last subsection. Note that the conclusion of Lemma 8.8 holds for this Ψ. Suppose that $F_{\mathbf{s}}$ belongs to $\mathcal{S}_2(n,l,\mathbf{j}) \cap Y_0(n,l,\mathbf{j},\zeta)$ for a parameter $\mathbf{s} \in R$. Then, by definition, there exists a point $w_1 \in \mathcal{C}(F)$ such that

$$d_{\mathbf{J}}(F_{\mathbf{s}}^n(\mathbf{j}), \mathbf{c}(w_1;F_{\mathbf{s}})) < 2B_g \exp(-\lambda^+(l)n(r-3)).$$

Since $\mathbf{c}(\cdot\,;F_{\mathbf{s}})=\mathbf{c}(\cdot\,;F)\colon \mathbf{B}(\mathcal{C}(F),\varrho_g)\to\mathbf{J}^{r-2}\Gamma$ is a C^1-mapping whose first-order differentials are bounded by some constant C_g, it follows that

$$\begin{aligned} d_{\mathbf{J}}(F^n_{\mathbf{s}}(\mathbf{j}),\mathbf{c}(F^n_{\mathbf{s}}(\mathbf{j})^{(0)};F_{\mathbf{s}})) &\leqslant d_{\mathbf{J}}(F^n_{\mathbf{s}}(\mathbf{j}),\mathbf{c}(w_1;F_{\mathbf{s}}))+d_{\mathbf{J}}(\mathbf{c}(w_1;F_{\mathbf{s}}),\mathbf{c}(F^n_{\mathbf{s}}(\mathbf{j})^{(0)};F_{\mathbf{s}}))\\ &<(1+C_g)d_{\mathbf{J}}(F^n_{\mathbf{s}}(\mathbf{j}),\mathbf{c}(w_1;F_{\mathbf{s}}))<C_g\delta_n^{\nu(r-3)}. \end{aligned}$$

Hence the image $\Psi(\mathbf{s})$ is contained in

$$\prod_{q=\nu+1}^{r-2}[-C_g\delta_n^{\nu(r-3)-(s-q)},C_g\delta_n^{\nu(r-3)-(s-q)}]\subset\mathbf{R}^{r-\nu-2}.$$

We arrive at the estimate

$$\mathbf{m}_{\mathbf{R}^{r-\nu-2}}(\{\mathbf{t}\in R\mid F_{\mathbf{t}}\in\mathcal{S}_2(n,l,\mathbf{j})\cap Y_0(n,l,\mathbf{j},\zeta)\})\leqslant C_g\prod_{q=\nu+1}^{r-2}\delta_n^{\nu(r-3)-(s-q)},$$

which holds uniformly for $F\in\mathcal{S}_2(n,l,\mathbf{j})\cap Y_0(n,l,\mathbf{j},\zeta)$, provided that n is larger than some constant C_g.

Now we apply Lemma 3.20. Fix a small number $0<T<1$ such that

$$\max_{|t_q|\leqslant T}\left\|\sum_{q=\nu+1}^{r-2}t_q\psi_{q,n,\zeta}\right\|_{C^s}\leqslant r\max_{\nu\leqslant q\leqslant r-2}\|\psi_q\|_{C^s}T<\varrho_s(d),$$

where $\varrho_s(d)$ is that in Lemma 3.18. Note that we can take T independently of n. Put $X=\mathcal{S}_2(n,l,\mathbf{j})\cap Y_0(n,l,\mathbf{j},\zeta)$ and $T_i=T$ in Lemma 3.20. Then the assumption (26) holds from the choice of T, and the subset Y in the statement of Lemma 3.20 is contained in $Y_1(n,l,\mathbf{j},\zeta)$ from the condition (V1), which we have proved above. Therefore we can obtain, as the conclusion,

$$\frac{\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_2(n,l,\mathbf{j})\cap Y_0(n,l,\mathbf{j},\zeta))\cap\mathbf{D}^r(d))}{\mathcal{M}_s(\Phi_G^{-1}(Y_1(n,l,\mathbf{j},\zeta)))}\leqslant C_gT^{-r+\nu+2}\prod_{q=\nu+1}^{r-2}\delta_n^{\nu(r-3)-(s-q)},$$

provided that n is larger than some constant C_g. Then the subsets $Y_0(n,l,\mathbf{j},\zeta)$ for $\zeta\in\mathbf{L}\big(\frac{1}{20}\delta_n\big)$ cover $C^r(M,M)$, while the intersection multiplicity of the subsets $Y_1(n,l,\mathbf{j},\zeta)$ for $\zeta\in\mathbf{L}\big(\frac{1}{20}\delta_n\big)$ is bounded by some absolute constant. Hence we can conclude that the measure $\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_2(n,l,\mathbf{j}))\cap\mathbf{D}^r(d))$ is bounded by

$$\begin{aligned} &C_gT^{-r+\nu+2}\prod_{q=\nu+1}^{r-2}\delta_n^{\nu(r-3)-(s-q)}\\ &\qquad=C_gT^{-r+\nu+2}\exp\left((r-\nu-2)\left(-(r-3)+\frac{2s-r-\nu+1}{2\nu}\right)\lambda^+(l)n\right). \end{aligned}$$

The subset $\mathcal{S}_2$ is contained in the closed subset

$$\mathcal{S}_2' := \bigcap_{n \geqslant B_g} \bigcup_{l=1}^{l_0} \bigcup_{\mathbf{j} \in \mathbf{Q}(n,l)} \mathcal{S}_2(n,l,\mathbf{j})$$

by Corollary 8.5. Hence the measure $\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_2') \cap \mathbf{D}^r(d))$ is bounded by

$$C_g T^{-r+\nu+2} \sum_{l=1}^{l_0} \#\mathbf{Q}(n,l) \exp\left((r-\nu-2)\left(-(r-3)+\frac{2s-r-\nu+1}{2\nu}\right)\lambda^+(l)n\right)$$

for sufficiently large n. From the estimate (98) on the cardinality of $\mathbf{Q}(n,l)$ and the condition in the choice of $\lambda^\pm(l)$, this converges to 0 exponentially fast as $n \to \infty$. Thus we conclude that $\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_2) \cap \mathbf{D}^r(d)) = \mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_2') \cap \mathbf{D}^r(d)) = 0$. As d is an arbitrary positive number, $\mathcal{M}_s(\Phi_G^{-1}(\mathcal{S}_2)) = 0$ or $\mathcal{S}_2$ is shy with respect to $\mathcal{M}_s$.

Suppose that $r \geqslant 19$. Then the inequality (3) holds for $s = r+3$ and $\nu = 3$, and so $\mathcal{M}_{r+3}(\Phi_G^{-1}(\mathcal{S}_2')) = 0$ for any $G \in C^r(M, \mathbf{T})$. This implies that $\mathcal{U} \setminus \mathcal{S}_2'$ is dense. Therefore $\mathcal{S}_2$ is contained in the closed nowhere dense subset $\mathcal{S}_2'$.

Appendix A. Proof of Corollary 2.3

To see that Corollary 2.3 follows from Theorem 2.2, it is enough to show the following result:

LEMMA A.1. *If X is a Borel subset in $C^r(M, \mathbf{T}^2)$ that is timid for the class $\mathcal{Q}_s^r$ of measures for some $s > r$, then the subset*

$$Y = \{F(z,t) \in C^r(M \times [-1,1]^k, \mathbf{T}) \mid \mathbf{m}_{\mathbf{R}^k}(\{t \in [-1,1]^k \mid F(\cdot, t) \in X\}) > 0\}$$

is timid for the class of Borel finite measures on $C^r(M \times [-1,1]^k, \mathbf{R}^2)$ that are quasi-invariant along $C^s(M \times [-1,1]^k, \mathbf{R}^2)$.

Proof. Take a mapping $G \in C^r(M \times [-1,1]^k, \mathbf{T})$ and put $G_0(z) = G(z,0)$. We define the mapping

$$P(f, \mathbf{t}) := G(\cdot, \mathbf{t}) - G_0(\cdot) + f(\cdot, \mathbf{t}) \colon C^r(M \times [-1,1]^k, \mathbf{R}^2) \times [-1,1]^k \longrightarrow C^r(M, \mathbf{R}^2),$$

so that

$$\Phi_{G_0} \circ P(f, \mathbf{t}) = G(\cdot, \mathbf{t}) + f(\cdot, \mathbf{t}).$$

Let $\mathcal{N}$ be a Borel finite measure on $C^r(M \times [-1,1]^k, \mathbf{R}^2)$ that is quasi-invariant along $C^s(M \times [-1,1]^k, \mathbf{R}^2)$. Then the measure $(\mathcal{N} \times \mathbf{m}_{\mathbf{R}^k}|_{[-1,1]^k}) \circ P^{-1}$ on $C^r(M, \mathbf{R}^2)$ belongs to $\mathcal{Q}_s^r$, and so we have $(\mathcal{N} \times \mathbf{m}_{\mathbf{R}^k})((\Phi_{G_0} \circ P)^{-1}(X)) = 0$ from the assumption. This and Fubini's theorem imply that $\mathcal{N} \circ \Phi_G^{-1}(Y) = 0$ and hence the claim of the lemma. □

Appendix B. Proof of Lemma 3.18

We use the definitions and results in the book [20] by Skorohod. We consider the functions $e_{nm}(x,y)=\exp\big(2\pi\sqrt{-1}\,(nx+my)\big)$ for $n,m\in\mathbf{Z}$ as a complete orthonormal basis of the space $L^2(\mathbf{T},\mathbf{m})$. Let $A\colon L^2(\mathbf{T},\mathbf{m})\to L^2(\mathbf{T},\mathbf{m})$ be the operator defined by

$$A\Bigg(\sum_{(n,m)\in\mathbf{Z}^2} a_{nm}e_{nm}\Bigg)=\sum_{(n,m)\in\mathbf{Z}^2}(n^2+m^2+1)^{-1/2}a_{nm}e_{nm}.$$

Let $\mathcal{N}$ be the Gaussian measure [20, §5] on $L^2(\mathbf{T},\mathbf{m})$ whose characteristic function is $\Theta(\psi)=\exp\big(-\frac{1}{2}(A^{2s-3}\psi,\psi)_{L^2}\big)$. Then $\mathcal{N}$ is supported on the Sobolev space $W^{s-3}:=A^{s-3}(L^2(\mathbf{T},\mathbf{m}))$. We can see, from [20, §16, Theorem 2], that $\mathcal{N}$ is quasi-invariant along $W^{s-3/2}\supset C^{s-1}(\mathbf{T},\mathbf{R})$ and that

$$\frac{d(\mathcal{N}\circ\tau_\psi^{-1})}{d\mathcal{N}}(\varphi)=\exp\big((A^{-s}\psi,A^{-s+3}\varphi)_{L^2}-\tfrac{1}{2}\|A^{-s+3/2}\psi\|_{L^2}^2\big)\leqslant\exp(\|\psi\|_{W^s}\|\varphi\|_{W^{s-3}})$$

for $\psi\in W^s$ and $\mathcal{N}$-almost every $\varphi\in W^{s-3}$.

We show that the measure $\mathcal{N}$ is actually supported on $C^{s-3}(\mathbf{T},\mathbf{R})$. We follow the argument in the proof of the fact that the measure corresponding to Brownian motion is supported on the space of continuous paths [11]. Let $\varphi^{(s-3)}$ be one of the $(s-3)$rd partial differentials of φ. Denoting the expectation with respect to the measure $\mathcal{N}$ by $E(\cdot)$, we have

$$E(|\varphi^{(s-3)}(z)-\varphi^{(s-3)}(w)|^5)\leqslant\text{const}\cdot d(w,z)^{5/2},$$

because the distribution of $\varphi^{(s-3)}(z)-\varphi^{(s-3)}(w)$ is a Gaussian distribution with average 0 and variance bounded by

$$\sum_{(n,m)\in\mathbf{Z}^2}(\min\{2,(n^2+m^2+1)^{1/2}d(z,w)\}(n^2+m^2+1)^{-3/4})^2\leqslant\text{const}\cdot d(z,w).$$

By the Borel–Cantelli lemma, there is a constant $i_0>0$ for $\mathcal{N}$-almost every φ such that

$$|\varphi^{(s-3)}(2^{-i}p,2^{-i}q)-\varphi^{(s-3)}(2^{-i}p',2^{-i}q')|^5\leqslant 2^{-i/3}$$

for $i>i_0$ and $p,q,p',q'\in\mathbf{Z}$ such that $|p-p'|\leqslant 1$ and $|q-q'|\leqslant 1$. This implies that $\varphi^{(s-3)}$ is continuous for $\mathcal{N}$-almost every φ, and hence $\mathcal{N}$ is supported on $C^{s-3}(\mathbf{T},\mathbf{R})$.

As $C^{s-3}(\mathbf{T},\mathbf{R}^2)$ is naturally identified with $C^{s-3}(\mathbf{T},\mathbf{R})\times C^{s-3}(\mathbf{T},\mathbf{R})$, we regard the product $\mathcal{N}\times\mathcal{N}$ as a measure on $C^{s-3}(\mathbf{T},\mathbf{R}^2)$. Put $\mathcal{M}_s=(\mathcal{N}\times\mathcal{N})\circ\pi^{-1}$, where $\pi\colon C^{s-3}(\mathbf{T},\mathbf{R}^2)\to C^{s-3}(M,\mathbf{R}^2)$ is the mapping that corresponds to the restriction to M. Then $\mathcal{M}_s$ satisfies the conditions in the lemma.

References

[1] ALVES, J. F., SRB measures for nonhyperbolic systems with multidimensional expansion. *Ann. Sci. École Norm. Sup.*, 33 (2000), 1–32.

[2] ALVES, J. F., BONATTI, C. & VIANA, M., SRB measures for partially hyperbolic systems whose central direction is mostly expanding. *Invent. Math.*, 140 (2000), 351–398.

[3] BONATTI, C. & VIANA, M., SRB measures for partially hyperbolic systems whose central direction is mostly contracting. *Israel J. Math.*, 115 (2000), 157–193.

[4] BUZZI, J., SESTER, O. & TSUJII, M., Weakly expanding skew-products of quadratic maps. *Ergodic Theory Dynam. Systems*, 23 (2003), 1401–1414.

[5] FALCONER, K. J., The Hausdorff dimension of self-affine fractals. *Math. Proc. Cambridge Philos. Soc.*, 103 (1988), 339–350.

[6] FELDMAN, J., Nonexistence of quasi-invariant measures on infinite-dimensional linear spaces. *Proc. Amer. Math. Soc.*, 17 (1966), 142–146.

[7] GOLUBITSKY, M. & GUILLEMIN, V., *Stable Mappings and Their Singularities*. Graduate Texts in Math., 14. Springer, New York–Heidelberg, 1973.

[8] HOFBAUER, F. & KELLER, G., Quadratic maps without asymptotic measure. *Comm. Math. Phys.*, 127 (1990), 319–337.

[9] HUNT, B. R., SAUER, T. & YORKE, J. A., Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc.*, 27 (1992), 217–238.

[10] — Prevalence: an addendum. *Bull. Amer. Math. Soc.*, 28 (1993), 306–307.

[11] IKEDA, N. & WATANABE, S., *Stochastic Differential Equations and Diffusion Processes*, 2nd edition. North-Holland Math. Library, 24. North-Holland, Amsterdam, 1989.

[12] KOZLOVSKI, O. S., Axiom A maps are dense in the space of unimodal maps in the C^k topology. *Ann. of Math.*, 157 (2003), 1–43.

[13] OSELEDETS, V. I., A multiplicative ergodic theorem. Characteristic Lyapunov exponents of dynamical systems. *Trudy Moskov. Mat. Obshch.*, 19 (1968), 179–210 (Russian); English translation in *Trans. Moscow Math. Soc.*, 19 (1968), 197–231.

[14] PALIS, J., A global view of dynamics and a conjecture on the denseness of finitude of attractors, in *Géométrie complexe et systèmes dynamiques* (Orsay, 1995). *Astérisque*, 261 (2000), 335–347.

[15] PERES, Y. & SOLOMYAK, B., Absolute continuity of Bernoulli convolutions, a simple proof. *Math. Res. Lett.*, 3 (1996), 231–239.

[16] PESIN, YA. B., Characteristic Lyapunov exponents and smooth ergodic theory. *Uspekhi Mat. Nauk*, 32 (1977), 55–112 (Russian). English translation in *Russian Math. Surveys*, 32 (1977), 55–114.

[17] POLLICOTT, M. & SIMON, K., The Hausdorff dimension of λ-expansions with deleted digits. *Trans. Amer. Math. Soc.*, 347 (1995), 967–983.

[18] PUGH, C. & SHUB, M., Ergodic attractors. *Trans. Amer. Math. Soc.*, 312 (1989), 1–54.

[19] SHEN, W., On the metric properties of multimodal interval maps and C^2 density of Axiom A. *Invent. Math.*, 156 (2004), 301–403.

[20] SKOROHOD, A. V., *Integration in Hilbert Space*. Ergeb. Math. Grenzgeb., 79. Springer, New York–Heidelberg, 1974.

[21] STINCHCOMBE, M. B., The gap between probability and prevalence: loneliness in vector spaces. *Proc. Amer. Math. Soc.*, 129 (2001), 451–457.

[22] TSUJII, M., A measure on the space of smooth mappings and dynamical system theory. *J. Math. Soc. Japan*, 44 (1992), 415–425.

[23] — Positive Lyapunov exponents in families of one-dimensional dynamical systems. *Invent. Math.*, 111 (1993), 113–137.

[24] — Fat solenoidal attractors. *Nonlinearity*, 14 (2001), 1011–1027.

[25] VIANA, M., Multidimensional nonhyperbolic attractors. *Inst. Hautes Études Sci. Publ. Math.*, 85 (1997), 63–96.

MASATO TSUJII
Department of Mathematics
Graduate School of Science
Hokkaido University
Sapporo, Hokkaido 060-0810
Japan

Received January 13, 2004