On Dynamical Systems.

Nicolai Theodorus Antonius Haydn.

Thesis submitted for the degree of Ph.D..
June 1986.

Mathmatics Institute,
University of Warwick.

Contents.

Acknowledgement.

I am grateful to Robert Mackay, who was my supervisor while writing the first part of this thesis, and I am grateful to Peter Walters for taking this position during the preparation of the rest of this thesis. I feel especially thankful for the interesting problems they selected and for uncountably many discussions.

For my final year I acknowledge financial support from the University of Warwick.

Declaration.

The content of this thesis is original, except where indicated otherwise.

The first part is submitted for publication in 'Communications in Mathematical Physics'. The second part is to appear in 'Ergodic theory and Dynamical Systems'.

Yet not your words only but mine own fantasy
That will receive no object for my head,
But ruminates on necromantic skill.
Philosophy is odious and obscure.
Both law and physic are for petty wits.
Divinity is basest of the three,
Unpleasant, harsh, contemptible and vile.
'Tis magic, magic that hath ravished me.
Then, gentle friends, aid me in this attempt,
And I, that have with subtle syllogisms
Gravelled the pastors of the English Church
And made the flowering pride of Worcestershire
Swarm to my problems as the infernal spirits
On sweet Musaeus when he came to hell,
Will be as cunning as Agrippa was,
Whose shadow made all Europe honour him

Christopher Marlow,
Doctor Faustus, Act one.

Summary.

Part A.

We prove existence of smooth invariant circles for area preserving twist maps close enough to integrable using renormalisation. The smoothness depends upon that of the map and the Liouville exponent of the rotation number.

Part B.

Ruelle and Capocaccia gave a new definition of Gibbs states on Smale spaces. Equilibrium states of suitable function there on are known to be Gibbs states. The converse in discussed in this paper, where the problem is reduced to shift spaces and there solved by constructing suitable conjugating homeomorphisms in order to verify the conditions for Gibbs states which Bowen gave for shift spaces, where the equivalence to equilibrium states is known.

Part C.

On subshifts which are derived from Markov partitions exists an equivalence relation which idendifies points that lie on the boundary set of the partition. In this paper we restrict to symbolic dynamics. We express the quotient space in terms of a non-transitive subshift of finite type, give a necessary and sufficient condition for the existence of a local product structure and evaluate the Zeta function of the quotient space. Finally we give an example where the quotient space is again a subshift of finite type.

Introduction.

This thesis consists of three parts, each of which is self-contained and has an abstract at the beginning of each part.

In the first part we consider the problem of invariant curves of cylinder maps. There exists already an impressive list of existence proofs of invariant curves of mappings of an annulus to itself assuming the map to be near enough to integrable. In all these proofs Newton's method of approximation is used at some point. The first part of this thesis focuses mainly on a method which enables us finally to construct invariant curves in an explicit way, which to some degree is suitable for numerical exploition. We use a renormalisation approach which was first introduced to study circle maps but has subsequently proved popular in the analysis of area-preserving maps of an annulus to itself. A more detailed description is to be found in the introcuction to part A itself.

The second part deals with Gibbs states in terms of a new definition which uses conjugating homeomorphisms. Let Ω be a compact metric space with metric $d(.,.)$. A map ψ from some open $U \subset \Omega$ into Ω is called conjugating, if $d(T^k \circ \psi(x), T^k(x)) \to 0$ for $|k| \to \infty$ uniformly in $x \in U$. Let F be a Hölder continuous real valued function on Ω and set

$$g(z) = \exp \sum_{k \in \mathbb{Z}} (F \circ T^k \circ \psi(z) - F \circ T^k(z)).$$

A probability measure ν is called a Gibbs state for F if

$$_U\int \tau \circ \psi \; g \; d\nu = {}_{\psi(U)}\int \tau \; d\nu,$$

holds true for all bounded and measurable functions $\tau\colon \psi(U) \to \mathbb{R}$ and all conjugating homeomorphisms $\psi\colon U \to \psi(U)$, for $U = U_\psi$ some open set in Ω.

Let (Ω,T) be a Smale space (a compact, metric space with an expanding homeomorphism and a local product structure), then for any continous function $F\colon \Omega \to \mathbb{R}$ the pressure $P(T,F)$ can be defined by the variational principle

$$P(T,F) = \sup_\rho (h_T(\rho) + {}_\Omega\int F \; d\rho),$$

where ρ runs over all T-invariant probability measures over Ω. Here $h_T(\rho)$ is the measure theoretic entropy with respect to T and ρ. A measure, for which the supremum is attained, is called an equilibrium state. Equilibrium states for F are also Gibbs states for F . This is proven in Ruelle's book [4] theorem 7,18. The converse, there referred to as an open question (cf. [4] p. 170), will be demonstrated in part B of this thesis.

There are several definitions of Gibbs states, for instance in terms of interactions, which was exploited especially by Ruelle in his famous book about thermodynamic formalism [4]. On the other hand, restricted to symbolic formalism, to which level we may ascend by introducing Markov partitions, we are given a powerful instrument to deal with the question of whether Gibbs states are necessarily equilibrium states. Our approach is to make the connection to the work of Bowen and all those who worked in his tradition.

For an nxn-matrix A of 0's and 1's we define the (one-sided) shiftspace

$$\Sigma_A{}^+ = \{x \in \prod_{0 \ldots +\infty} \{1,\ldots,n\}: \; A[x_i, x_{i+1}] = 1 \;\; \forall \; i \in \mathbb{Z}^+\}.$$

For a $\lambda \in (0,1)$ we can define a metric on $\Sigma_A{}^+$ by $d(x,y) = \lambda^n$, where $n = n(x,y) = \max\{m: \; x_i = y_i \;\; \forall \;\; 0 \le i \le m\}$. The continuous map $\sigma\colon \Sigma_A{}^+ \to \Sigma_A{}^+$ given by $(\sigma x)_i = x_{i+1}$, $i \ge 0$, is alled the (one-sided) shift. Note, that σ is a bounded to one local homeomorphism. A continuous function $f\colon \Sigma_A{}^+ \to \mathbb{R}$ is said to be exponentially decreasing (or hölder continuous) if $|f(x) - f(y)| \le C\alpha^n$ for constants $C > 0$ and $\alpha \in (0,1)$, where $d(x,y) \le \lambda^n$. The Ruelle operator $\mathfrak{L}_f$ maps exponentially decreasing functions again into such functions and is defined by $(\mathfrak{L}_f g)(x) = \sum_{\sigma y = x} g(y) e^{f(y)}$.

Ruelle's Perron-Frobenius Theorem: (See [1]) Suppose $\Sigma_A{}^+$ is topologically mixing (e.g. $A^n > 0$ for n large enough) and f is hölder continuous. Then there are $\theta > 0$, a positive and continuous function h on $\Sigma_A{}^+$ and a measure ν on $\Sigma_A{}^+$ such that $\mathfrak{L}_f h = \theta h$, $\mathfrak{L}_f{}^* \nu = \theta \nu$, $\nu(h) = 1$ and $\lim_{m \to \infty} \|\theta^{-m} \mathfrak{L}_f{}^m g - \nu(g) h\| = 0$ for all continuous $g\colon \Sigma_A{}^+ \to \mathbb{R}$.

Then $\mu = h\nu$ is a probability measure on $\Sigma_A{}^+$: $\mu(g) = \nu(hg) =$

$\int h(x)g(x)d\nu$ for all measurable $g: \Sigma_A^+ \to \mathbb{R}$, and is invariant under the shift: $\mu(g) = \mu(g \circ \sigma)$. Furthermore, for a positive constant c and a real P, μ satisfies the following inequalities:

(*) $$e^{-c} \leq \exp(mP - \sum_{0 \leq k < m} f \circ \sigma^k(x))\mu(U(x_0 \ldots x_m)) \leq e^c,$$

for all $x \in \Sigma_A^+$ and $m \in \mathbb{N}$, where $U(x_0 \ldots x_m)$ is the cylinder set $\{y \in \Sigma_A^+ : y_i = x_i$ for all $0 \leq i \leq m\}$. The number P is called the pressure of f. Probability measures satisfying these inequalities are said to be Gibbs states. With the help of Ruelle's Perron-Frobenius Theorem it is shown that for every hölder continuous f there is a unique Gibbs state. In fact, this result holds only in the case of one dimensional lattice systems.

Let Q be a disjoint cover of Σ_A^+ by unions cylinder sets, and set $h(\nu, Q) = -\sum_{q \in Q} \nu(q) \log \nu(q)$, then $h(\nu) = \sup_Q h(\nu, Q)$ is the measure theoretical entropy of ν, and $h(\Sigma_A) = \sup_\nu h(\nu)$ is the topological entropy of Σ_A^+ which coincides with the maximal (positive) eigenvalue of the transition matrix A.

A measure that achieves according to the variational principle the suppremum of the following expression

$$P(f) = \sup_\nu \{h(\nu) + \int f d\nu\},$$

where ν runs over all probability measures on Σ_A^+, is called an equilibrium state.

Theorem: (see [1]) Given that $f: \Sigma_A^+ \to \mathbb{R}$ is hölder continuous, then there is a unique equilibrium state of f. This also satisfies the Gibbs condition (*).

This theorem characterises Gibbs states (*) as equilibrium states. To get an equilibrium state it is therefore enough to verify (*). This is what will be done in Part B. There it will turn out to be convenient to use the two-sided shift instead of the one-sided one introduced here. A hölder continuous function F defined on the two-sided shift

$$\Sigma_A = \{x \in \prod_{-\infty \ldots +\infty} \{1, \ldots, n\} : A[x_i, x_{i+1}] = 1 \ \forall \ i \in \mathbb{Z}\}$$

is always cohomologous to a function f depending only on the positive coordinates, i.e. a function on Σ_A^+. This means, there is a continuous $u: \Sigma_A \to$

R (depending on F) so that $F = f + u \circ \sigma - u$. Gibbs and equilibrium states are not affected by adding a coboundary $u \circ \sigma - u$, they are the same for F as for f.

In the last part we are concerned about a special kind of equivalence relations which occour in symbolic dynamics. The questions treated there arise from Markov partitions of Axiom A diffeomorphisms (cf. Bowen [2]). For a small enough partition one gets a shiftspace and a projection onto the original manifold where the diffeomorphism acts conjugate to the shift. It is known that a subshift of finite type can be isomorphic only to an Axiom A diffeomorphism over a non-wandering set of zero dimension. It is therefore clear that the boundary set, i.e. the set of points which have a pre-image in the shift consisting of more than one point preserves the essential structure of the non-wandering set, despite the fact that it has measure zero for any smooth measure.

We begin part C by demonstrating that it is enough to consider strings of some certain length whenever we want to decide whether a relation induces an equivalence relation on Σ_A. In the following three sections we restrict to equivalence relations that have finite equivalence classes. In that case the quotient space can be described by means of a non-transitive subshift, which has a partial ordering. Maximal elements with respect to this ordering correspond to points in the quotient space. This formulation will be used in section 3 to express the topology on the quotient space in terms of cylinder sets of this new shift space. In the same section we give a necessary and sufficient condition on the existence of a local product structure on the quotient space. In section 4 we evaluate the Zeta function under the assumption made that the equivalence classes are finite. It turns out, that in this more general context the Zeta function is given by Mannings product formula (see [3]). Finally, in the last section we investigate a special kind of shift spaces, which have quotient spaces that are again subshifts of finite type.

References.

[1] R. Bowen; *Equilibrium, states and the ergodic theory of anosov diffeomorphism.* SLN # 470, (1975).

[2] R. Bowen; *On Axiom A diffeomorphisms.* AMS Regional Conf. Proc., 35 (1978).

[3] A. Manning; Axiom A diffeomorphism have rational Zeta function. LMS Bulletin, 3 (1971), pp 215 - 220.

[4] D. Ruelle; *Thermodynamic Formalism;* Addison Wesley, (1978).

On invariant curves under renormalisation.

Nicolai T. A. Haydn
February 1985, revised November 1986.
Mathematics Institute, University of Warwick, Coventry CV4 7AL, U. K..

Abstract: We prove existence of smooth invariant circles for area preserving twist maps close enough to integrable using renormalisation. The smoothness depends upon that of the map and the Liouville exponent of the rotation number.

1. Introduction.

The existence or not of invariant circles in area preserving maps is of great importance since it relates directly to problems of stability and confinement. The first results on this problem are due to Kolmogorov, Arnold [1] and Moser [6] assuming that the map has a certain twist property, which roughly means that points on the cylinder which lie on a higher level than others mover faster. Arnold proved existence of analytic invariant circles for nearly integrable analytic maps. Moser proved existence of invariant circles if the map is C^p-close enough to integrable, the number p of derivatives required depending quadratically on the number of derivatives required for the circle. For example, he proved existence of C^1 circles for C^{333} maps close enough to integrable. Later on the number of derivatives sufficient for C^1 curves was lowered to $p > 3$, by Rüssmann [8]. In the general case one considers irrational rotation numbers ω, which saftisfy a diophantine condition $|\omega - p/q| \geq Cq^{-\gamma-2}$ for a positive C, where γ is called the Liouville exponent. Rüssmann also got a better relationship between the class C^p of the map and the Liouville exponent γ of the rotation number of the circle; he showed that $p > 2\gamma + 3$ is sufficient for existence of a continuous circle, for maps C^p-close enough to integrable. Restricting to numbers of constant type $(\gamma=0)$ Herman proved existence of invariant C^s-circles for $p > s + 1$, $p > 3$ [3].

In this paper a similar result to Rüssmann's will be proven by a different method, which gives us also differentiability of the invariant curves, the number of derivatives depending linearly on p and γ.

In the second section of this paper we present the main result and introduce the renormalisation operator acting on commuting pairs of twist maps. The third section treats commutativity and provides a method to get estimates on the derivatives of the generating functions if bounds on the highest order derivatives are known. In the fourth part we proof convergence of the renormalisation. Finally, in the fifth section the invariant curves will be constructed, first only with Lipshitz continuity which is achieved by 'pulling back' a single point on the invariant circle with an increasing number of iterations, a method introduced by Rand [7]. Then, in proposition 16 the smoothness result is proven by employing a similar procedure which 'pulls back' smooth curves. Finally, in lemma 17 the method of proposition 16 is copied and used to show that the invariant circle may be parametrised in some smooth manner.

The renormalisation introduced in section 2 can be extended to a more general class of mappings of the plane into itself. However, in the case where the rotation number is the golden mean $(1+\sqrt{5})/2$, MacKay has proved the convergence of the renormalisation operator by explicitly evaluating its eigenvalues and eigendirctions. It then turns out that the eigenvalues in modulus are all strictly less than one, so long as one restricts to area preserving twist maps.

I am grateful to R. S. MacKay, for without his encouragement and advice this paper would not exist and to D. A. Rand for reading the manuscript.

2. Renormalisation.

Let φ be an area preserving twist map: φ: $S^1\times\mathbb{R}$ ↺. $(x,y)\to(x',y') = \varphi(x,y)$ with x an angle variable and dx'/dy bounded away from zero, either positive or negative. Area preservingness of φ is reflected in the fact that the Jacobian $D\varphi$ is one. Instead of φ we will often consider a lift Φ: $\mathbb{R}^2$ ↺, $(x,y)\to(x',y') = \Phi(x,y)$, where now x is a variable with domain $\mathbb{R}$, and Φ is periodic: $x'(x-1,y) = x'(x,y) - 1$. Alternatively, if we set for the shift by one in the x-direction R: $(x,y) \to (x-1,y)$, the periodicity reads $\Phi\circ R = R\circ\Phi$. Denote by π_x the projection onto the x-axis. If for a point $\xi\in\mathbb{R}^2$ the limit $\omega = \lim_{q\to\infty}\pi_x\Phi^q(\xi)/q$ exists, then it is called the rotation number of ξ. This is the same as there exist $p = p(q)$ such that

$$\pi_x\Phi^q R^p\xi = q\omega - p + o(q) \quad \text{as } q\to\infty.$$

If ξ belongs to a circle on which φ is topologically conjugate to a rotation, then we have the stronger statement

$$\pi_x\Phi^{q[n]}R^{p[n]}\xi \to 0 \quad \Leftrightarrow \quad p[n]\omega - q[n]\to 0 \tag{2-1}$$

as n goes to infinity, for sequences $\{(p[n],q[n])\in\mathbb{N}^2,\ n\in\mathbb{N}\}$. Choose an irrational number ω. There is no loss of generality assuming that ω lies in the unit interval $(0,1)$ since the rotation numbers of a point under different lifts coincide modulo 1. Let

$$\omega = [m[0],m[1],\ldots] = (m[0]+(m[1]+(\ldots)^{-1})^{-1})^{-1} \in \mathbb{R}/\mathbb{Z}$$

be the continued fraction expansion of ω and let $p[n]/q[n]$ be the convergents found in the well-known way by setting $p[0] = 0$, $p[1] = q[0] = 1$, $q[1] = m[0]$ and by using the recursion formulas

$$p[i+1] = m[i]p[i] + p[i-1], \quad q[i+1] = m[i]q[i] + q[i-1]$$

for $i \in \mathbb{N}$. From the construction of the convergents of ω it seems natural to introduce a method which in a similar way generates inductively the expressions $\Phi^{q[n]}R^{p[n]}$ that appear in (2-1). But first let us generalise the notion of rotation number to pairs of commuting, area preserving twist maps (U,T), where U and T for the moment are commuting maps of $\mathbb{R}^2$ into itself. We say the point $\xi \in \mathbb{R}^2$ has rotation number ω, if for all sequences $\{(p[n],q[n]) \in \mathbb{N}^2, n \in \mathbb{N}\}$ one has

$$\pi_x U^{q[n]}T^{p[n]}/\max(p[n],q[n]) \to 0 \quad \Leftrightarrow \quad p[n]/q[n] \to \omega$$

for $n \to \infty$.

For any natural number m the renormalisation operator N_m is defined acting on pairs of commuting, area preserving twist maps by N_m: $(U,T) \to (\Lambda T \Lambda^{-1}, \Lambda T^m U \Lambda^{-1})$, for a suitable coordinate transform Λ. Later on we shall be more precise about this point, for the time being we remark only that Λ depends on (U,T) and the number m. And indeed, for iterates of the renormalisation operator we obtain the same expressions as they appear in (2-1):

$$N_{m[n]}\cdots N_{m[0]}(U,T) = (\Lambda_0\cdots\Lambda_n U^{q[n-1]}T^{p[n-1]}\Lambda_n^{-1}\cdots\Lambda_0^{-1}, \Lambda_0\cdots\Lambda_n U^{q[n]}T^{p[n]}\Lambda_n^{-1}\cdots\Lambda_0^{-1}).$$

Take an oriented homotopically non-trivial curve $\mathfrak{C}$ on $S^1\times\mathbb{R}$. Then $\varphi(\mathfrak{C})$ is again a homotopically non trivial curve on $S^1\times\mathbb{R}$ and $\mathfrak{C}$ and $\varphi(\mathfrak{C})$ enclose some area. We count the portion of this area which lies to the right (as determined by the orintation) of $\mathfrak{C}$ positive, the one to the left negative and call the difference the flux of φ. The flux is independent of the choice of $\mathfrak{C}$ and is sometimes called the Calabi invariant. If the flux is non-zero, then φ shifts on average along the cylinder. In that case of course one cannot expect that there is any φ-invariant homotopically non-trivial curve in $S^1\times\mathbb{R}$.

Define for $\gamma \geq 0$: $\mathfrak{L}(\gamma) = \{\omega \in \mathbb{R}\backslash\mathbb{Q}: \;\; \exists\, C > 0,$ such that for all $q \in \mathbb{N}, \;\; p \in \mathbb{Z}, \;\; q|\omega q - p| \geq Cq^{-\gamma}\}$. For a given ω the number γ is called Liouville exponent and C the Liouville constant. Naturally, all γ that are bigger than some Liouville exponent are again Liouville exponents of ω. However, we cannot define the Liouville exponent by the infinimum over all γ. For instance the numbers of type Roth $\bigcap_{\gamma>0}\mathfrak{L}(\gamma)$ have full Lebesgue measure, but $\mathbb{L}(0)$, the numbers of constant type (the m[n] in the continued fraction expansion are bounded) have Lebesgue measure zero.

We say μ is a transitive invariant circle of φ if

(i) μ is invariant under φ, and

(ii) μ is on $S^1 \times \mathbb{R}$ represented by a homotopically non-trivial circle.

We say φ is of class $C^{j+\text{Lipshitz}}$ if φ is j-times differentiable and the j-th derivative is Lipshitz (of course all the lower ones as well). Let $\mathfrak{z} = (1 + \sqrt{5})/2 = 1 + [1,1,1,...]$ be the golden mean and set $\beta = \log(1+\mathfrak{z}^{-2})/\log \mathfrak{z}$. The main result we shall prove in this paper is the next theorem.

Theorem 1. Let $\varphi: S^1 \times \mathbb{R} \circlearrowleft$ be an area preserving twist map of class C^L, $L \geq 4$, with zero flux, and let $\omega \in \mathbb{L}(\gamma)$, suppose $\gamma < \gamma^*(L)$, where $\gamma^*(L) = (L-6-\beta)/4 + \sqrt{[(L-2-\beta)^2/16 + 1/2]}$.
Then if φ is in the C^L-topology close enough to the affine shear $(x,y) \to (x+\omega+y,y)$ in a neighbourhood of $\{(x,y) \in S^1 \times \mathbb{R}:\ y = 0\}$ (depending on γ, L and the Liouville constant C), then

(i) there exists a transitive and invariant circle of rotation number ω and class $C^{j+\text{Lipshitz}}$ for $j \leq \psi = \psi(\gamma) = \max(0, [L-3-\beta-\gamma]) \in \mathbb{N}$.

(ii) restricted to this invariant circle, φ acts $C^{j+\text{Lipshitz}}$-conjugate to a rotation by ω on S^1 for all $j \leq \rho$.

Observe that the upper bound for the Liouville exponent $\gamma^*(L)$ in particular is strictly less than $(L-3-\beta)/2$.

We begin with a definition.

Definition 2: A pair of twist mappings (U,T) near enough to affine shearings (near enough so that all objects we are talking about are defined), $U,T: \mathfrak{B} \to \mathbb{R}^2$ for a region $\mathfrak{B} \subset \mathbb{R}^2$, is **normalised** if

(i) $U(Y) \cap Y = (0,1)$, where $Y = \mathfrak{B} \cap \{(x,y) \in \mathbb{R}^2:\ x = 0\}$,

(ii) $T(Y) \cap Y = (0,0)$,

(iii) $U(Y)$ has at the point $(0,1)$ slope one.

(iv) U and T commute in some sense to be described later on.

Finally, in the proof of the theorem, we shall set $(U,T) = (\Lambda\Phi\Lambda^{-1}, \Lambda\Phi^{m[0]}R\Lambda^{-1})$, where Λ is an affine coordinate transformation designed to normalise the pair $(\Phi, \Phi^{m[0]}R)$. Of particular interest are the normalised affine shearings. They form a one parameter family which we call the **simple line**

$$(U_\infty:\ (x,y) \to (x+y-1,y),\ T_{\infty,c}:\ (x,y) \to (x+cy,y)),$$

where c is a real, positive parameter. In the course of the proof of the theorem it turns that c will always be bigger than some number arbitrary close but less than 1, depending how near Φ is to affine.

Let R be the right halfplane $\{(x,y) \in \mathbb{R}^2\colon\ x \geq 0\}$, L be the left halfplane $\{(x,y) \in \mathbb{R}^2\colon\ x < 0\}$ and define

$$\Psi(\xi) = \begin{cases} U(\xi) & \text{if } \xi \in D \cap R \\ T(\xi) & \text{if } \xi \in D \cap L. \end{cases}$$

The map Ψ is not continuous at Y, but if we identify $U(Y)$ with $T(Y)$ by identifying $U(\xi)$ with $T(\xi)$ for $\xi \in Y$, Ψ can be considered as a continuous map of a cylinder $\mathbf{Z}$ to itself. We will need this fact only to define the flux for a pair of mappings and will not any more refer to it later on. The flux of the pair (U,T) is defined as the flux of Ψ as a cylinder map.

We say μ to be a transitive invariant circle of (U,T) if

(i) μ is invariant under Ψ,

(ii) μ is on $\mathbf{Z}$ represented by a homotopically non-trivial circle.

The renormalisation operator N_m, $m \in \mathbb{N}$, maps normalised maps to normalised maps, and is defined as $N_m\colon\ (U,T) \to (\Lambda T \Lambda^{-1}, \Lambda T^m U \Lambda^{-1})$, with the coordinate transform Λ to get a new normalised pair with R and L interchanged. A priori Λ may be any coordinate change. We will use coordinate transformations that are close to affine (see also MacKay [4]). The transformations will be introduced in lemma 10, where we show that the renormalisation operator is a contraction in the neighbourhood of the simple line. In lemma 9 the non linear part h of Λ brings $\Lambda U \Lambda^{-1}$ close to the affine shear U_∞.

Given a point $\xi \in R$, then a natural number m will be associated to it in the following way: m is determined so that $U\xi \in L, TU\xi \in L, \ldots, T^m U\xi \in L$ and $T^{m+1}U\xi \in R$, provided $U\xi \in L$ all iterates are defined. The coordinate transform Λ normalises the pair $(T, T^m U)$, and since $\Lambda T^m U\xi \in R$ we are the position to determine a new number m'. Iterating this process yields a sequence $\{m[n] \in \mathbb{N}\colon\ n \in \mathbb{N} \cup \{0\}\}$ which in general will break off after finitly many steps, as $\xi_n = \Lambda_n \Lambda_{n-1} \cdots \Lambda_0 \xi$ for some n leaves the domain of $U_n = \Pi_1(N_{m[n]} \cdots N_{m[0]}(U,T))$, where $\Pi_1(U,T) = U$, if any of the iterates $T_n{}^k U_n \xi_n$ is no longer defined, or if $\pi_x T_n U_n \xi$ and $\pi_x U_n \xi$ are no longer negative. If the sequence does not break off, then $\omega = [m[0], m[1], \ldots]$ is the rotation number of ξ. We note here, that later on we will use as well the notation $\Pi_2(U,T) = T$ and $T_n = \Pi_2(N_{m[n]} \cdots N_{m[0]}(U,T))$.

We fix an irrational number ω and then we search for a point on Y which has ω as rotation number. This is in effect the reverse of the procedure desrcibed in the previous paragraph, where we fixed ξ and then went on to determine its rotation number. The operator $N_{m[0]}$ sends the rotation number ω to $\omega' = \omega^{-1} - m[0] = [m[1], m[2], \ldots] \in \mathbb{R}/\mathbb{Z}$, $N_{m[0]}$ acts as a shift on the components of the continued fraction expansion.

It is convenient to decompose N_m. Define

$$N^\wedge:\ (U,T) \to (TU,T),$$
$$N^*:\ (U,T) \to (\Lambda T \Lambda^{-1}, \Lambda U \Lambda^{-1}).$$

Hence $N_m = N^* \circ N^{\wedge m}$. Write for further use $(U[p],T[p]) = N^{\wedge p}(U,T)$. This decomposition plays an essential role in lemma 8. The same applies to section five, where the iterates $(U[p],T[p])$ are needed to construct approximations of the invariant curve out of small segments.

Renormalisation by using affine coordinate transforms leaves the simple line invariant, and N_m acts on the parameter by $c \to m+c^{-1}$. It is not hard to see, that any affine and area preserving map T^*: $\mathbb{R}^2 \to \mathbb{R}^2$ which commutes with U_∞ is in fact a shearing parallel to the y-axis, i.e. is of the form $(x,y) \to (x+cy+a, y)$, where a and c are some constants. We shall need this fact in section 3. Will return to it shortly in the context of generating functions.

Now switch to generating functions (ν,τ) (cf. Mather [5]). An area preserving (or any other measure preserving) twist map $U: (x,y) \to (x',y')$ can always be represented by a generating function $\nu(x,x')$: $\mathbb{R}^2 \to \mathbb{R}$ (here the twist property is important, U need not necessarily be near an affine shear in some other sense). The arguments x and x' are the x-components of a point (x,y) and its image $(x',y') = U(x,y)$. The y-components are obtained by differentiating as follows

$$\partial_x \nu(x,x') = -y,$$
$$\partial_{x'} \nu(x,x') = y'.$$

If U is in C^l then ν is in C^{l+1}. Similarly, T can be expressed by a generating function $\tau(x,x')$: $\mathbb{R}^2 \to \mathbb{R}$.The decomposition of the renormalisation operator reads now

$$N^\wedge:\ (\nu,\tau) \to (\nu \oplus \tau, \tau),$$
$$N^*:\ (\nu,\tau) \to (\Lambda \tau \Lambda^{-1}, \Lambda \nu \Lambda^{-1}),$$

where $(\nu \oplus \tau)(x,x') = \nu(x,x'') + \tau(x'',x')$ for an intermediate coordinate x'' chosen to satisfy the stationarity condition

$$\partial_{x''}\big(\nu(x,x'') + \tau(x'',x')\big) = 0,$$

where $\partial_{x''}$ stands for d/dx''. The stationarity condition say in particular that there is a unique intermediate y-coordinate $y'' = \partial_{x''}\nu(x,x'') = -\,\partial_{x''}\tau(x'',x)$. In our case x'' will always be defined unambiguously, because the second derivative with respect to x'' will be negative and bounded away from zero for (ν,τ) close enough to the simple line.

Let $c > 0$ some number, and define in the x,x'-plane the rectangles

$$\mathfrak{E} = \{(x,x') \in \mathbb{R}^2:\ 1/4 \le x-x' \le 3,\ |x+x'| \le 5\},$$
$$\mathfrak{F} = \{(x,x') \in \mathbb{R}^2:\ \max(-2\sqrt{\tfrac{1}{2}}, -3c/4) \le x-x' \le \min(2\sqrt{\tfrac{1}{2}}, 2c),\ |x+x'| \le 4\}.$$

Throughout the paper, (ν,τ) will be defined in $(\mathfrak{E},\mathfrak{F})$ and ν, τ will commute in $\mathfrak{F}$. Let $(\nu_\infty,\tau_{\infty,c})$ be defined in $(\mathfrak{E},\mathfrak{F})$, then U_∞: $(x,y) \to (x+y-1, y)$, is defined in

$$\mathfrak{E}^* = \{(x,y) \in \mathbb{R}^2:\ -3 \le y \le 3/4,\ |x - (1-y)/2| \le 5/2\},$$

and $T_{\infty,c}$: $(x,y) \to (x+cy,y)$ is defined in

$$\mathfrak{F}^* = \{(x,y) \in \mathbb{R}^2:\ \max(-2\sqrt{\tfrac{1}{2}}/c, -2) \le y \le \min(3/4, 2\sqrt{\tfrac{1}{2}}/c),\ |x - y/2c| \le 2\}.$$

In particular $\mathfrak{F}^* \subset \mathfrak{E}^*$, and $\mathfrak{F}$ is the same for $c \ge \sqrt{\tfrac{1}{2}}8/3$, which in fact is the more interesting portion of values c may take. In (x,x')-coordinates the normalising conditions get more concrete:

(i) $\partial_{x'}\nu(x,x')|_{(0,0)} = 1$,

(ii) $\partial_{x'}\tau(x,x')|_{(0,0)} = 0$,

(iii) $\partial_x\partial_{x'}\nu(x,x')|_{(0,0)} = -1$, and

(iv) commutativity in $\mathfrak{F}$, i.e. $(\nu\oplus\tau)(x,x') = (\tau\oplus\nu)(x,x')$ for $(x,x') \in \mathfrak{F}$.

The last condition (for (ν,τ) close enough to the simple line), $\nu(x,x'') + \tau(x'',x') = \tau(x,\hat{x}'') + \nu(\hat{x}'',x')$, involves in general two different intermediate coordinates x'' and $\hat{x}''$. In generating functions the simple line is represented by

$$(\nu_\infty(x,x') = ((x-x'-1)^2/2,\ \tau_{\infty,c}(x,x') = (x-x')^2/2c)), \quad c > 0.$$

Let $\tau(x,x') = \mathfrak{a}x + \mathfrak{b}x' + \mathfrak{c}x^2 + 2\mathfrak{d}xx' + \mathfrak{e}x'^2$ be a map that containes no higher than quadratic terms and commutes with ν_∞, where $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}, \mathfrak{e}$ are some constants. According to a remark made above, we have necessarily $\mathfrak{c} = -\mathfrak{d} = \mathfrak{e}$. We shall use this fact in the proof of proposition 5.

For convenience we write $(\nu[u],\tau[u]) = N^{\wedge u}(\nu,\tau)$, $N^{\wedge 0} = \mathrm{id}$, for $u \in \mathbb{N}\cup\{0\}$. At this point we will outline in more detail how the renormalisation works using affine shearings. Given an irrational $\omega = [m[0],m[1],\ldots]$ and set ν^* and τ^* for the generating functions associated with U_0 and T_0. Let

$$(\nu,\tau) = N_{m[n-1]}N_{m[n-2]}\cdots N_{m[0]}(\nu^*,\tau^*).$$

(There is no index on the ν's and τ's to denote the iteration under the renormalisation operator.) In order to obtain $(\nu',\tau') = N_{m[n]}(\nu,\tau)$ we have to apply $N^\wedge$ $m = m[n]$ - times before interchanging ν and τ and rescaling to get finally (ν',τ'). Say, $c = c[n-1]$ is the parameter value after $n-1$ renormalisations. Then

$$(\nu,\tau) = (\nu_\infty,\tau_{\infty,c}) = \left((x-x'-1)^2/2,\ (x-x')^2/2c\right).$$

Once applying $N^\wedge$ we obtain

$$(\nu[1],\tau[1]) = \left((x-x'-1)^2/[2(1+c)],\ (x-x')^2/2c\right),$$

and for arbitrary $u \in \mathbf{N}$ in general

$$(\nu[u],\tau[u]) = \left((x-x'-1)^2/[2(1+uc)],\ (x-x')^2/2c\right).$$

In particular we read off the rescaling factors which determine the linear part of the coordinate transform Λ. In the y-direction one must stretch by a factor $-(1+mc)$ and in x-direction by $-c/(1+mc) = -1/(m+c^{-1})$. Set $c = c[n] = \Delta_{n-1}^{-1}$ and let $c[0]$ be the initial parameter value, then one finds

$$c[n+1]^{-1} = \Delta_n = [m[n],m[n-1],\ldots,c[0]] \in \mathbf{R}/\mathbf{Z},$$

and in particular $c[n] = m[n-1] + \Delta_{n-2}$. The Jacobian of the coordinate transformation Λ_n turns out to be

$$D\Lambda_n = \begin{vmatrix} -\Delta_n^{-1} & 0 \\ 0 & -\Delta_n^{-1}\Delta_{n-1}^{-1} \end{vmatrix}.$$

Additional to that rescaling, Λ_n shifts in y-direction by 1.

Derivatives with respect to the first, respectively second variable are denoted by ∂_1 and ∂_2 and the symbol ∂ itself stands for any derivative. For w in some interval $I \subset \mathbf{R}$, with $[0,1] \subset I$, write $D_w = \partial_1 + w\partial_2$. In lemma 8 we shall get more precise about the range of w. Finally, before we restate theorem 1 in the context of pairs of commuting twist mappings we introduce some notation which will in effect not be used until section 4. We do it here because for reading section 3 it may be helpful to know that the sizes η and Δ which appear there are the same as those defined here. Define

$$\varepsilon_{p,u} = \max_{w\in I,\, j+k=p} \left(\|\partial_2^j D_w^k(\nu[u]-\nu_\infty)\|_{\mathfrak{C}},\ \|\partial_2^j D_w^k(\tau[u]-\tau_{\infty,c})\|_{\mathfrak{f}}\right),$$

for $1 \le p \le l = L+1$, $\|\,.\,\|$ is the supremum norm is as indicated on $\mathfrak{C}$ respectively $\mathfrak{f}$. The parameter value c is well chosen, for instance so that $\partial_1\partial_2(\tau[u]-\tau_{\infty,c})|_{(0,0)} = 0$. Let $\tau_{i,j} = \partial_1^i\partial_2^j\tau$, then the derivatives in $\{\tau_{i,j};\ i,j \ge$

$0,\ i+j = p\}$ ($\partial^p\tau$ denotes any one of them) are certainly contained in $\{\partial_2{}^j D_w{}^k\tau$: $j+k = p$ and $w \in [0,1]\}$ for $1 \le p \le l$ since D_0 is just ∂_1. Nevertheless, in lemma 8 it will turn out to be more convenient to work with D_w, ∂_2 instead of with ∂_1, ∂_2. Take as abbrevation $\eta_p = \varepsilon_{p,0}$, and to account for the special role the highest derivative l will take, we set $\eta(n) = \eta_l(n)$ which is understood to be η_l after n renormalisations. The main result of section 3, proposition 5, will show how the lower order derivatives can be estimated by the higher ones, i.e $\eta_p(n) \le \mathfrak{r}_p\eta_l(n)$, $p = 1,\dots,l-1$, for some constants $\mathfrak{r}_p$.

Theorem 3. Let (U,T) be a (commuting) pair of normalised maps in C^L, $L \ge 4$, with zero flux and suppose $\mathfrak{R}$ defined in some bounded region $\mathfrak{B} \subset \mathbb{R}^2$. Let (ν,τ) be their generating functions in C^l, $l = L+1$, and let $\omega \in \mathfrak{L}(\gamma)$ for $\gamma < \gamma^*(L)$ an irrational number in the unit interval. If we assume (ν,τ) in the C^l-topology on $(\mathfrak{E},\mathfrak{F})$ is near enough to the simple line, then

(i) the sequence of renormalised pairs of (ν,τ) converges in the C^l-topology on $(\mathfrak{E},\mathfrak{F})$ to the simple line of affine shearings,

(ii) the pair (U,T) possesses a transitive and invariant curve μ of class $C^{j+\text{Lipshitz}}$ for all $j \le \psi(\gamma)$, and

(iii) (U,T) acting on μ is $C^{j+\text{Lipshitz}}$-conjugate to $(\mathfrak{R}_{-1},\mathfrak{R}_\omega)$ on the x-axis for $j \le \psi$, where $\mathfrak{R}_\omega$: $x \to x+\omega \bmod 1$.

3. Commutativity.

The next lemma generalises the method of comparing coefficients to the case of Taylor expansions with remainder.

Lemma 4. Let θ be a positive number and let $P(x) = \sum_{0\le i\le n} a_i x^i$ be a polynomial of degree n with $|P(x)| \le \theta$, for x in an interval I centred at 0. Then there are numbers d_i, depending on the degree n of P and on the interval I, such that $|a_i| \le d_i\theta$ for $i = 0,\dots,n$.

Proof. Set without loss of generality $I = [-1,1]$ and decompose P into Chebycheff polynomials $P(x) = \sum_{0\le i\le n} c_i T_i(x)$. The $T_i x)$ are orthonormal with respect to a weighted integral. Hence

$$|c_i| = \left|\pi^{-1}2^{2i-2}\,{}_{-1}\!\int^{+1} P(x)T_i(x)(1-x^2)^{-1/2}dx\right|$$

$$\le \pi^{-1}2^{2i-2}\sqrt{\left[{}_{-\pi}\!\int^{+\pi}|P(\cos\psi)|^2d\psi\right]}\cdot\sqrt{\left[{}_{-\pi}\!\int^{+\pi}(T_i(\cos\psi))^2d\psi\right]}$$

$$\le 2^{2i-2}\theta 2^{3/2-i} = 2^{i-1/2}\theta.$$

The coefficient of the highest power of $T_i(x)$ is 1, so we conclude $|a_n| \leq 2^{n-1/2}\theta$. Now replace $P(x)$ by $\sum_{0 \leq i \leq n-1} a_i x^i$ and θ by $\theta(1 + 2^{n-1/2})$. Repeating the argument proves the lemma. □

In this section η and Δ denote some (rather small) positive numbers. In particular one has to think of η so small that the product $\eta\Delta^{2-l}$ is (uniformly) bounded by some small constant. Whenever we write $O(\Delta)$ or $O(\eta)$ the indicated estimates are understood to be uniformly in Δ respectively η. This section aims to prove proposition 5, i.e. given bounds on some higher derivatives the point is then to estimate the lower ones.

Proposition 5. Let Δ and η be some small, positive numbers and suppose

(i) ν and τ commute in $\mathfrak{F}$;

(ii) $\|\partial^p \nu\|_{\mathfrak{F}} = O(\eta)$ for $p = 2,\dots,l$;

(iii) $\|\partial^l \tau\|_{\mathfrak{F}} = O(\eta)$;

(iv) $\|\partial^p(\tau - \tau_{\infty,c})\|_{\mathfrak{F}} = O(\Delta)$ for $p = 2,\dots,l-1$

(v) c is a parameter value satisfying $|\log \Delta c| \leq 2$, say.

Then there exist positive numbers $\mathfrak{t}_p$, $p = 1,\dots,l-1$, which are independent of Δ and η, so that $\|\partial^p(\tau - \tau_{\infty,c})\|_{\mathfrak{F}} \leq \mathfrak{t}_p\eta$.

The constants $\mathfrak{t}_p$ depend on p, where $\mathfrak{t}_p > \mathfrak{t}_{p+1}$ for $1 \leq p < l$, and they furthermore depend on l. In fact, $\mathfrak{t}_p$ increases when l increses and correspondingly the estimates on the derivatives have to be 'pulled back' a longer way.

In lemma 6 we shall use commutativity to improve the a priori bounds on $\partial^p(\tau - \tau_{\infty,c})$ by a factor Δ. Repeatedly applying lemma 6 finally will prove the claim made in proposition 5 for at least third order derivatives of τ. This procedure is often referred to as 'bootstrap'. At this point we use the commutativity assumption. We expand ν and τ into Taylor polynomials and compaire their coefficients. This provides $l(l+1)/2$ inequalities for the same number of coefficients in the Taylor expansion of τ. The remainder in the Taylor expansion is by hypothesis (ii) estimated in terms of $O(\eta)$, and the coefficients in the Taylor expansion of $\tau - \tau_{\infty,c}$ which are of lower than l-th order, will turn out to be of the same size $O(\eta)$. In fact, in lemma 6 we show smallness of higher than second order derivatives and the second order derivatives themselves will be dealt with in the proof of the proposition at the end of this section.

Lemma 6. Let $q, \Delta, \eta > 0$ real numbers, satisfying $\Delta^{[q-l]}\eta = O(\Delta)$, (where $[z] = \min(0,z)$) and suppose

(i) ν and τ commute in $\mathfrak{F}$;

(ii) $|\partial^p(\nu-\nu_\infty)|_{\mathfrak{C}} = O(\eta)$ for $p = 2,\dots,l$;
(iii) $|\partial^l\tau|_{\mathfrak{f}} = O(\eta)$;
(iv) $|\partial^p\tau|_{\mathfrak{f}} = O(\Delta^{[q-l]}\eta)$, for $p = 3,\dots,l-1$;
(v) $|\partial^2\tau|_{\mathfrak{f}} = O(\Delta)$.
Then $\qquad |\partial^p\tau|_{\mathfrak{f}} = O(\Delta^{[q+1-l]}\eta)$, for $p = 3,\dots,l-1$.

Proof. The second order derivatives remain untouched throughout the proof. They only have to be of size Δ which is exactly the factor by which hypothesis (iv) gets improved. We consider the commutativity condition

(3-1) $$\nu\bullet\tau - \tau\bullet\nu = \nu(x,x'') + \tau(x'',x') - \tau(x,\tilde{x}'') - \nu(\tilde{x}'',x') = 0,$$

where the arguments have to lie in $\mathfrak{C}, \mathfrak{f}, \mathfrak{f}, \mathfrak{C}$ respectively. There are two intermediate coordianates, x'' refers to $\nu\bullet\tau$ and $\tilde{x}''$ to $\tau\bullet\nu$. The stationarity conditions

$$\partial_{x''}[\nu(x,x'') + \tau(x'',x')] = 0,$$
$$\partial_{\tilde{x}''}[\tau(x,\tilde{x}'') + \nu(\tilde{x}'',x')] = 0$$

allows us to consider, for instance, x and $\tilde{x}''$ as variables and x'' and x' depending on them. The stationarity conditions for a pair of normalised generating functions $(\nu_\infty,\tau_{\infty,c})$, where $c \geq 1$ is some parameter, are

$$-x + x''(1+c^{-1}) - x'c^{-1} + 1 = 0$$
$$-xc^{-1} + \tilde{x}''(1+c^{-1}) - x' - 1 = 0.$$

We see that $x \sim x''$ and $\tilde{x}'' \sim x'$ and one sees that $(x,\tilde{x}'')$ is a good choice for coordinates (the other one is (x'',x'), we shall return to it in the second part of the proof). More precisely, $x, \tilde{x}''$ can vary over some interval which uniformly in c is boundet from below (and above) so that all arguments that involve x'',x' will not leave their domain. We expand (3-1) into a Taylor polynomial around some $x_0, x''_0, \tilde{x}''_0, x'_0$ and denote for the sake of clarity the differences again by $x, x'', \tilde{x}'', x'$. The foot points $x_0, x''_0, \tilde{x}''_0, x'_0$ may be chosen so that $x'_0 = 0$ and $\nu_{01}(\tilde{x}''_0,x'_0) = 0$. Since ν is up to an error of $O(\eta)$ near the simple line we have $\tilde{x}''_0 = 1 + O(\eta)$ and therefore $-\nu_{10}(\tilde{x}''_0,x'_0) = O(\eta)$. Furthermore $y' = \tau_{01}(x''_0,x'_0) = \nu_{01}(\tilde{x}''_0,x'_0) = 0$ and therefore $x''_0 = 0$ because of the normalistion condition (iii) $\tau_{01}|_{(0,0)} = 0$.

(i) Differentiate (3-1) with respect to x, hence

$$d_x(\nu\bullet\tau - \tau\bullet\nu) = \nu_{10}(x,x'') - \tau_{10}(x,\tilde{x}'') + [\tau_{01}(x'',x') - \nu_{01}(\tilde{x}'',x')]d_x x'.$$

We expand each term into a Taylor series and obtain using hypothesis (ii), (iii) that

$$\sum_{(i,j)\geq(1,0),l>i+j} \{x^{i-1}[\nu_{i,j}x^{"j} - \tau_{i,j}x^{\sim"j}] + x'^{i-1}[\tau_{j,i}x^{"j} - \nu_{j,i}x^{\sim"j}]d_x x'\}/(j!(i-1)!) = O(\eta)$$

(3-2)

for $(x_0+x,x^")$, $(x_0+x,x_0^{\sim"}+x^{\sim"})$, etc. lying in $\mathfrak{C}$, $\mathfrak{f}$, etc.. Let us prune away all those terms that are small by assumption (ii):

$$\sum_{(i,j)\geq(1,0),l-1>i+j\geq2} \{-x^i x^{\sim"j}\tau_{i+1,j} + x'^i x^{"j}\tau_{j,i+1}(d_x x')\}/(j!i!) - \tau_{10}(x_0,x^{\sim"}_0)$$

$$+ \nu_{10}(x_0,x^"_0) - x^{\sim"}\tau_{11} + x(1-\tau_{20}) - x^" + (x^{\sim"}+x^"\tau_{11}-x'(1-\tau_{02}))d_x x' = O(\eta),$$

(3-2*)

since $\nu_{01}(x^{\sim"}_0,x'_0) = \tau_{01}(x^"_0,x'_0) = 0$. Here we used that, up to an error of $O(\eta)$, $\nu_{11} = -1$, $\nu_{20} = +1$, $\nu_{02} = +1$ and $\nu_{i,j} = O(\eta)$ for $i+j \geq 3$. The term $x^"$ causes the biggest trouble (more than all the rest together).

(A) We expand $x^"$ into a Taylor series (remember $x^"_0 = 0$):

$$x^" = \sum_{(s,t)\geq(0,0),l\geq s+t\geq1} (d_{x^{\sim"}}{}^s d_x{}^t x^")x^{\sim"s}x^t/(s!t!),$$

where d denotes total derivatives. We have to compute the derivatives $d_{x^{\sim"}}{}^s d_x{}^t x^"|(x_0,x^{\sim"}_0)$. The implicit function theorem applied to the first of the two stationarity conditions yields

$$\partial_x x^" = -\nu_{11}/(\nu_{02} + \tau_{20}) = (1+\tau_{20})^{-1} + O(\eta),$$

$$\partial_{x'} x^" = -\tau_{11}/(\nu_{02} + \tau_{20}) = -\tau_{11}(x^"_0,x'_0)/(1+\tau_{20}) + O(\Delta\eta) = O(\Delta),$$

where the ν's are evaluated at $(x_0,x^"_0)$ and the τ's at $(x^"_0,x'_0)$. Our aim is to take the estimates on the derivatives of τ as they are given by hypothesis (iv) and to improve them by a factor Δ. Set $V_3 = \max_{3\leq u<l}(\|\partial^u\tau\| + \|\partial^u\nu\|)$ which is of order $O(\eta\Delta^{[q-l]})$. Whenever possible we shall therefore ignore terms of order $O(\Delta V_3)$. (Observe that $V_3 = O(\Delta)$, i.e. $V_3{}^2 = O(\Delta V_3)$) Firstly, we deduce

(3-3) $$\partial_{x'}{}^s x^" = -\tau_{1,s}/(\nu_{02} + \tau_{20}) + O(\Delta V_3) = -\tau_{1,s}/(1+\tau_{20}) + O(\Delta V_3),$$

for $s \geq 1$. Here we used that the remainder $O(\Delta V_3)$ involves products of at least third derivatives of ν and τ together with at least second derivatives of τ which are of size Δ by hypothesis (iv). Using once more the implicit function theorem we derive from the second of the stationarity conditions:

$$d_{x^{\sim"}} x' = \partial_{x^{\sim"}} x' = -[\tau_{02}(x_0,x^{\sim"}_0) + \nu_{20}(x^{\sim"}_0,x'_0)]/\nu_{11} = 1 + \tau_{02} + O(V_3),$$

$$d_x x' = \partial_x x' = -\tau_{11}(x_0,x^{\sim"}_0)/\nu_{11}(x^{\sim"}_0,x'_0) = \tau_{11} + O(\eta),$$

since, due to hypothesis (ii), $\nu_{11}(x^{\sim}{}_0,x'_0)$ is -1. up to an error of size $O(\eta)$. Next we shall show that $d_x{}^t x" = \partial_x{}^t x" + O(\Delta V_3)$ for $t \geq 2$. Clearly $d_x x" = \partial_x x" + (d_x x')\partial_{x'} x"$, and differentiating once with respect to x provides $d_x{}^2 x" = \partial_x{}^2 x" + O(\Delta V_3)$, since products of the form $(d^2 x')(\partial_{x'} x")$ are of size $O(\Delta V_3)$, where one of the two factors is differentiated at least twice while the other one at least once. For some $t \geq 3$ let us assume $d_x{}^{t-1} x" = \partial_x{}^{t-1} x" + O(\Delta V_3)$ holds true. Now we increase the number of derivatives by one, hence

$$d_x{}^t x" = \partial_x{}^t x" + (d_x x')\partial_{x'}\partial_x{}^{t-1} x" + O(\Delta V_3) = \partial_x{}^t x" + O(\Delta V_3)$$

if t is at least 2. Let t be 2 and it follows for the partial derivative

$$\partial_x{}^2 x" = -\partial_x[\nu_{11}/(\nu_{02}+\tau_{20})] + O(\Delta V_3)$$

$$= -\nu_{21}(\nu_{02}+\tau_{20})^{-1} + \nu_{11}\nu_{12}(\nu_{02}+\tau_{20})^{-2} + O(\Delta V_3) = O(\Delta V_3).$$

The same holds true for higher derivatives, i.e. $d_x{}^t x" = O(\Delta V_3)$ for $t \geq 2$ $(s = 0)$, because derivatives of the remainder are again of size $O(\Delta V_3)$. For $(s,t) \geq (1,2)$ we claim

$$d_{x^{\sim}}{}^s d_x{}^t x" = (d_{x^{\sim}} x')^s \partial_{x'}{}^s \partial_x{}^t x" + O(\Delta V_3).$$

As we have seen, this is true for $s = 0$. Fix t and let us assume the formula is proven for $s-1$ for some $s \geq 1$. Differentiating once with respect to $x^{\sim}$ yields

$$d_{x^{\sim}}{}^s d_x{}^t x" = (d_{x^{\sim}} x')^s \partial_{x'}{}^s \partial_x{}^t x" + (s-1)(d_{x^{\sim}} x')^{s-2}(d_{x^{\sim}}{}^2 x')\partial_{x'}{}^{s-1}\partial_x{}^t x" + O(\Delta V_3)$$

$$= (d_{x^{\sim}} x')^s \partial_{x'}{}^s \partial_x{}^t x" + O(\Delta V_3).$$

Furtherore, since $d_{x^{\sim}} x'$ is $1+c^{-1}+O(\Delta V_3)$ and is therefore uniformly bounded, and since the partial derivatives $\partial_x{}^t x"$ are of size $O(\Delta V_3)$ for $t \geq 2$, we obtain

$$d_{x^{\sim}}{}^s d_x{}^t x" = O(\Delta V_3),$$

for $(s,t) \geq (0,2)$. In the cases where $s \geq 1$ and $t = 1$ it remains some extra work to do. We begin with $d_x x" = \partial_x x" + (d_x x')\partial_{x'} x"$ and differentiate it once with respect to $x^{\sim}$:

$$d_{x^{\sim}} d_x x" = (d_{x^{\sim}} x')\partial_{x'}\partial_x x" + (d_{x^{\sim}} d_x x')\partial_{x'} x" + (d_{x^{\sim}} x')(d_x x')\partial_{x'}{}^2 x" = O(\Delta V_3).$$

The same estimate holds true for higher $x^{\sim}$-derivatives. We summarize:

$$d_{x^{\wedge\prime\prime}}{}^{s} d_x{}^{t} x'' = O(\Delta V_3)$$

for all $(s,t) \geq (0,1)$, $(s,t) \neq (0,1)$. Finally we have to treat the case $s \geq 0$ and $t = 0$. Indeed $d_{x^{\wedge\prime\prime}} x'' = (d_{x^{\wedge\prime\prime}} x')\partial_{x'} x''$, and let us assume that $d_{x^{\wedge\prime\prime}}{}^{s-1} x'' = (d_{x^{\wedge\prime\prime}} x')^{s-1} \partial_{x'}{}^{s-1} x'' + O(\Delta V_3)$ for some $s \geq 2$. By way of differentiating once we obtain

$$d_{x^{\wedge\prime\prime}}{}^{s} x'' = (d_{x^{\wedge\prime\prime}} x')^{s} \partial_{x'}{}^{s} x'' + (s-1)(d_{x^{\wedge\prime\prime}} x')^{s-2}(d_{x^{\wedge\prime\prime}}{}^{2} x')\partial_{x'}{}^{s-1} x'' + O(\Delta V_3)$$

$$= (d_{x^{\wedge\prime\prime}} x')^{s} \partial_{x'}{}^{s} x'' + O(\Delta V_3).$$

To finish off the last step, we set $\wp = (1+\tau_{20})$ and use (3-3) in conjunction with $d_{x^{\wedge\prime\prime}} x' = \wp + O(\Delta V_3)$, which leads us to

$$d_{x^{\wedge\prime\prime}}{}^{s} x'' = -\wp^{s-1} \tau_{1,s}(x''_0, x'_0) + O(\Delta V_3)$$

for $s \geq 2$. Apart from $d_x{}^{t} x''$ these are the only non-trivial derivatives (a priori not of size $O(\Delta V_3)$). In the case $s = 1$ up to a constant the same result holds true as we shall see in the next line:

$$d_{x^{\wedge\prime\prime}} x'' = [\tau_{02}(x_0, x^{\wedge\prime\prime}_0) + \nu_{20}(x^{\wedge\prime\prime}_0, x'_0)]\tau_{11}(x''_0, x'_0)/(1+\tau_{20}(x''_0, x'_0)) + O(\Delta V_3)$$

$$= -\wp^{*} \tau_{11}(x''_0, x'_0) + O(\Delta V_3),$$

where $\wp^{*} = (1+\tau_{02})/(1+\tau_{20}) \sim 1$. This completes the list of all derivatives of x''. So we end up with a polynomial in x and $x^{\wedge\prime\prime}$:

$$x'' = x[\wp^{-1} - \tau_{11}\tau_{11}(1+c^{-1})^{-1}] - \wp^{*} x^{\wedge\prime\prime} \tau_{11} - \sum_{2 \leq j} x^{\wedge\prime\prime j} \wp^{j-1} \tau_{1,j} + O(\Delta V_3),$$

(with the convention $\wp^{0} = \wp^{*}$) and in particular $x'' = x + O(\Delta)$.

(B) We expand x' into a Taylor series. Concerning the first order derivatives we know already

$$d_{x^{\wedge\prime\prime}} x' = -[\tau_{02}(x_0, x^{\wedge\prime\prime}_0) + \nu_{20}(x^{\wedge\prime\prime}_0, x'_0)]/\nu_{11}(x^{\wedge\prime\prime}_0, x'_0) = 1 + \tau_{02}(x_0, x^{\wedge\prime\prime}_0) + O(\eta)$$

$$d_x x' = -\tau_{11}(x_0, x^{\wedge\prime\prime}_0)/\nu_{11} = \tau_{11}(x_0, x^{\wedge\prime\prime}_0) + O(\Delta\eta).$$

For the higher order derivatives $d_{x^{\wedge\prime\prime}}{}^{s} d_x{}^{t} x'$ we derive; first for $s = 0$; $t \geq 1$ by differentiating it follows $d_x{}^{t} x' = \tau_{t,1}(x_0, x^{\wedge\prime\prime}_0) + O(\eta)$ for $t = 1, \ldots, l-1$. Furthermore we derive immediatly by differentiating with respect to $x^{\wedge\prime\prime}$: $d_{x^{\wedge\prime\prime}}{}^{s} d_x{}^{t} x' = \tau_{t,s+1}(x_0, x^{\wedge\prime\prime}_0) + O(\eta)$ for $(s,t) \geq (0,1)$. We are missing the pure $x^{\wedge\prime\prime}$

derivatives. Firstly

$$d_{x^{\sim}}{}^{2}x' = \tau_{03}(x_0,x^{\sim}{}_0)+\nu_{30}(x^{\sim}{}_0,x'_0)+\nu_{21}(x^{\sim}{}_0,x'_0)d_x x' + O(\eta)$$

$$= \tau_{03}(x_0,x^{\sim}{}_0) + O(\Delta V_3),$$

since $\nu_{30} = O(\eta)$ and $d_x x'$ is of order $O(\Delta)$. Similarly for higher derivatives: $d_{x^{\sim}}{}^{s}x' = \tau_{0,s+1} + O(\Delta V_3)$. Summarizing:

$$d_{x^{\sim}}{}^{s}d_x{}^{t}x' = \tau_{t,s+1} + O(\Delta V_3)$$

for $(s,t) \geq (0,0)$, $(s,t) \neq (0,0), (1,0)$. Finally, since $x'_0 = 0$, we get

$$x' = x\tau_{11} + x^{\sim}(1+\tau_{02}) + \sum_{(s,t)\geq(0,0),l\geq s+t\geq 2} x^{\sim s}x^{t}\tau_{t,s+1}/((s!t!) + O(\Delta V_3),$$

where the sum containes only summands of size $O(V_3)$. In particular $x' = x^{\sim} + O(\Delta)$ and $d_x x' = O(\Delta)$, as we would have expected from the affine case. Let us return to (3-2*) and insert the expressions for x'', x', in particular we use the fact $x''^{j}\tau_{j,1} = x^{j}\tau_{j,1}+O(\Delta V_3)$:

$$\sum_{(i,j)\geq(1,0),l-1>|(i,j)|\geq 2} -x^{i}x^{\sim j}\tau_{i+1,j}/(j!i!) - x''$$

$$- x^{\sim}\tau_{11} + x(1-\tau_{20}) + (x^{\sim}+x''\tau_{11}-x'(1-\tau_{02}))d_x x' + \nu_{10}(x_0,x_0'') = O(\Delta V_3)..$$

Hence

$$\sum_{(i,j)\geq(0,0),l-1>i+j\geq 2} -x^{i}x^{\sim j}\tau_{i+1,j}/(j!i!) + \sum_{j\geq 2} x^{\sim j}{}^{j-1}\tau_{1,j}$$

$$+ \text{first order terms in } x \text{ and } x^{\sim} = O(\Delta V_3).$$

Lemma 4 once applied to this polynomial in x yields

$$\sum_{l-i-1>j\geq 0} -x^{\sim j}\tau_{i+1,j}/j! = O(\Delta V_3)$$

for $i = 2,\ldots,l-1$. A second time and we obtain

$$\tau_{i+1,j} = O(\Delta V_3)$$

for $(i,j) \geq (2,0)$. Since we differentialted the commutativity condition with respect to x we do not get the $\tau_{0,j}$, $j = 3,\ldots,l-1$. As it turns out, the summand x'' in (3-2*) spoils the $\tau_{1,j}$ for $j = 2,\ldots,l-2$, since they cancel out (up to a factor Δ).

(ii) The second part of this proof is devoted to $\tau_{i,j}$, $i = 0,1$, $j = 3-i,\ldots,l-i-1$.

We proceed similarly as in part (i), this time chosing x'',x' as coordinates. We differentiate (3-1) with respect to x' and obtain

$$d_{x'}(\nu\circledast\tau - \tau\circledast\nu) = [\nu_{10}(x,x'') - \tau_{10}(x,x^{\hat{}''})]d_{x'}x + \tau_{01}(x'',x') - \nu_{01}(x^{\hat{}''},x').$$

It follows by hypothesis (ii), (iii) that

$$\sum_{(i,j)\geq(1,0),l>|(i,j)|\geq 1} \{x^{i-1}[\nu_{i,j}x''^{j} - \tau_{i,j}x^{\hat{}''j}]d_{x'}x + x'^{i-1}[\tau_{j,i}x''^{j} - \nu_{j,i}x^{\hat{}''j}]\}/(j!(i-1)!) = O(\eta)$$

(3-4)

for (x_0+x,x''), $(x_0+x,x_0^{\hat{}''}+x^{\hat{}''})$, etc. in $\mathfrak{C}, \mathfrak{f}$, etc.. Let us prune away all those terms that are small by assumption (ii) and the first part of this proof:

$$\sum_{l-1>j\geq 2}[-x^{\hat{}''j}\tau_{1,j}(d_{x'}x) + x'^{j}\tau_{0,j+1}] + \sum_{l-2>j\geq 1} x'^{j}x''\tau_{1,j+1} + x^{\hat{}''} - \tau_{10}(x_0,x^{\hat{}''}_0)$$

(3-4*) $$+ \nu_{10}(x_0,x''_0) - (x^{\hat{}''}\tau_{11}+x(1-\tau_{20})-x'')(d_{x'}x) + x''\tau_{11} - x'(1-\tau_{02}) = O(\eta),$$

since $\nu_{01}(x^{\hat{}''}_0,x'_0) = \tau_{01}(x''_0,x'_0) = 0$. Here we used that, up to an error of $O(\eta)$, $\nu_{11} = -1$, $\nu_{20} = +1$, $\nu_{02} = +1$ and $\nu_{i,j} = O(\eta)$ for $i+j \geq 3$.

(A) Similar to (i) part (A), expand $x^{\hat{}''}$ into a Taylor polynomial. With the assistance of the implicit function theorem we deduce from the stationarity conditions

$$d_{x''}x = \partial_{x''}x = -(\nu_{02} + \tau_{20})/\nu_{11} = 1 + \tau_{20} + O(\eta),$$

$$d_{x'}x = \partial_{x'}x = -\tau_{11}/\nu_{11} = \tau_{11} + O(\eta),$$

where the ν's are evaluated at (x_0,x''_0) and the τ's at (x''_0,x'_0). Furthermore, exploiting the stationarity condition that determines $x^{\hat{}''}$ we obtain

$$\partial_{x'}x^{\hat{}''} = -\nu_{11}/(\tau_{02} + \nu_{20}) = (1+\tau_{02})^{-1} + O(\eta),$$

$$\partial_{x}x^{\hat{}''} = -\tau_{11}/(\tau_{02} + \nu_{20}) = -(1+\tau_{02})^{-1}\tau_{11} + O(\Delta\eta),$$

and this time the τ's are evaluated at $(x_0,x^{\hat{}''}_0)$ and the ν's at $(x^{\hat{}''}_0,x'_0)$. We write the arguments only where necessary. Combining the last equations we get

$$d_{x'}x^{\hat{}''} = \partial_{x'}x^{\hat{}''} + (d_{x'}x)\partial_{x}x^{\hat{}''}$$

$$= (1+\tau_{02})^{-1} - (1+\tau_{02})^{-1}\tau_{11}(x''_0,x'_0)\tau_{11}(x_0,x^{\hat{}''}_0) + O(\Delta V_3),$$

$$d_{x''}x^{\hat{}''} = (d_{x''}x)\partial_{x}x^{\hat{}''} = -\tau_{11}(x_0,x^{\hat{}''}_0) + O(\Delta V_3).$$

Differentiating with respect to x'' yields

$$d_{x''}{}^2 x^{\hat{}''} = -(d_{x''}x)\tau_{21} + O(\Delta V_3) = O(\Delta V_3),$$

because $\tau_{21} = O(\Delta V_3)$ as a result of the first part. Hence $d_{x''}{}^s x^{\hat{}''} = O(\Delta V_3)$ for all $s \geq 2$. From (3-4*) it is clear that the only derivatives of $x^{\hat{}''}$ that are important are $d_{x''}{}^s d_{x'}{}^t x^{\hat{}''}$ for $s = 0,1$, $t = 1,\dots,l-s-1$. Firstly for $s = 0$, we obtain

$$d_{x'}{}^2 x^{\hat{}''} = d_{x'}\{(1+\tau_{02})^{-1} - (1+\tau_{02})^{-1}\tau_{11}(x''_0,x'_0)\tau_{11}(x_0,x^{\hat{}''}_0) + O(\Delta V_3)\}$$

$$= (1+\tau_{02})^{-2}(d_{x'}x)\tau_{12} - (1+\tau_{02})^{-1}[\tau_{12} + (d_{x'}x)\tau_{21}] + O(\Delta V_3),$$

$$= -(1+\tau_{02})^{-1}\tau_{12} + O(\Delta V_3),$$

since $(d_{x'}x)$ is of size $O(\Delta)$. As already mentioned earlier, the remainder $O(\Delta V_3)$ is a sum of products which consists of at least two factors of at least third derivatives of ν and at least second derivatives of τ. This means in practice that we neglect derivatives of the remainder. Thus

$$d_{x'}{}^t x^{\hat{}''} = -q\tau_{1,t}(x''_0,x'_0) + O(\Delta V_3),$$

for $t \geq 2$, where $q = (1+\tau_{02})^{-1}$. By the same argument we get

$$d_{x''}d_{x'}{}^t x^{\hat{}''} = -q\tau_{2,t}(x''_0,x'_0) + O(\Delta V_3) = O(\Delta V_3)$$

for $t \geq 1$, because $\tau_{2,t} = O(\Delta V_3)$. It follows readily that $d_{x''}{}^s d_{x'}{}^t x^{\hat{}''} = O(\Delta V_3)$ for all $(s,t) \geq (1,1)$. Finally we get for $x^{\hat{}''}$ the following expression

$$x^{\hat{}''} = x^{\hat{}''}_0 + qx'[1 - \tau_{11}\tau_{11}] - \sum_{2\leq j} x'^j q\tau_{1,j} - x''\tau_{11} + O(\Delta V_3).$$

(B) Additionally to $x^{\hat{}''}$ we shall need the Taylor polynomial of $x = x(x'',x')$. We know from (ii) part (A) that $d_{x'}x = \tau_{11}(x''_0,x'_0) + O(\eta)$ and therefore $d_{x''}{}^s d_{x'}{}^t x = \tau_{1+s,t}(x''_0,x'_0) + O(\eta)$ for all $(s,t) \geq (0,1)$. We are missing the pure x''-derivatives. The first was computed above:

$$d_{x''}x = [\nu_{02}(x_0,x''_0) + \tau_{20}(x''_0,x'_0)] + O(\eta),$$

and differentiating yields

$$d_{x''}{}^2 = [\nu_{03}(x_0,x''_0) + \tau_{30}(x''_0,x'_0)] + (d_{x''}x)\nu_{12}(x_0,x''_0) + O(\eta) = O(\Delta V_3),$$

since the ν_{03}, $\nu_{12} = O(\eta)$ and $\tau_{30} = O(\Delta V_3)$ as a result of the first part. We have therefore $d_{x'}{}^t x = \tau_{1,t} + O(\Delta V_3)$, $t = 1,\dots,l-1$, and $d_{x''}{}^s d_{x'}{}^t x = O(\Delta V_3)$ for $s \geq 1$. We summarize

$$x = x_0 + x'\tau_{11} + x''(1+\tau_{20}) + \sum_{l-1>j\geq 2} x'^j\tau_{1,j} + O(\Delta V_3).$$

We now return to (3-4*) and use that $(d_x x) = O(\Delta)$, thus

$$\sum_{l-1>j\geq 2} x'^j[\tau_{0,j+1} - q\tau_{1,j}]/j! + x''\sum_{l-2>j\geq 1} x'^j\tau_{1,j+1} + x^{\wedge\prime\prime}{}_0 - \tau_{10}(x_0,x^{\wedge\prime\prime}{}_0) + \nu_{10}(x_0,x''_0)$$

$$+ qx'[1-\tau_{11}\tau_{11}] - (x^{\wedge\prime\prime}\tau_{11}+x(1-\tau_{20})-x'')(d_x x) - x'(1-\tau_{02}) = O(\Delta V_3).$$

Lemma 4 applied to this polynomial in x' leads to

(i) $\quad O + x''\tau_{12} = O(\Delta V_3), \qquad j = 1,$
(ii) $\quad \tau_{0,j+1} - q\,\tau_{1,j} + \tau_{1,j+1}x'' = O(\Delta V_3), \qquad j = 2,\dots,l-1,$

where O is some irrelevant constant $(= q[1 - \tau_{11}\tau_{11}][1 + \tau_{11}\tau_{11}] - 1 + \tau_{02})$. The term $i = 0$ does not make sense. Inspecting the linear terms in x'' we get $\tau_{1,i} = O(\Delta V_3)$ for $i = 2,\dots,l-1$. The constant terms then obviously lead to $\tau_{0,i} = O(\Delta V_3)$ for $i = 3,\dots,l-1$, since $q \sim 1$. This concludes the proof the lemma 6.

□

Proof of proposition 5. We begin with some q^* which satisfies condition (iv) of lemma 6, i.e. $\Delta^{[q^*-1]}\eta = O(\Delta)$. Applying lemma 6 improves q to $q+1$ until ΔV_3 is of size η (recall $V_3 = \max_{3\leq u<l}(\|\partial^u\tau\| + \|\partial^u\nu\|) = O(\eta\Delta^{[q-1]})$). ($q$ is not necessarily an integer, and in lemma 6 we ought to have written $\Delta V_3 + \eta$ instead of ΔV_3 to prevent that ΔV_3 may get smaller than η.) Hence $\|\partial^p\tau\| = O(\eta)$ $p = 3,\dots,l$, or $\|\partial^p\tau\| \leq \epsilon_p\eta$ for some numbers ϵ_p, and τ is up to an error of size η a polynomial of degree two. Let $\tau(x,x') = x^2\tau_{20}/2 + xx'\tau_{11} + x'^2\tau_{02}/2 + x\tau_{10} + x'\tau_{01} + \text{const.}$. By the remark made in section 2 it follows that $\tau_{20} = -\tau_{11} = \tau_{02}$ up to an error of size $O(\eta)$. The parameter c is chosen so that $(\tau-\tau_{\infty,c})_{11}|_{(0,0)} = 0$ and therefore $\|\partial^2(\tau-\tau_{\infty,c})\| = O(\eta)$. This concludes the proof of proposition 5, i.e. $\|\partial^p(\tau-\tau_{\infty,c})\| = O(\eta)$ for $p = 2,\dots,l$. □

4. Convergence proof.

Proposition 7: Let $\omega = [m[0],m[1],\dots]$ be an irrational number and let (ν,τ) be a pair normalised generating function. Set $\Delta_n = [m[n],m[n-1],\dots,c[0]]$ and $M_n = \prod_{0\leq i<n}\Delta_i^{-1}$, where $c[0]$ is determined so that $\partial_1\partial_2(\tau-\tau_{\infty,c[0]})|_{(0,0)} = 0$. Then there exist constants $\Theta > 0$, $\Xi > 0$ (depending on $m[0],c[0]$), and a sequence of coordinate changes $\{\Lambda_n;\ n \in \mathbb{N}_0\}$ that are near to affine, such that the renormalisation sequence of (ν,τ) converges in the C^l-topology on $(\mathbb{C},\mathcal{F})$

to the simple line, provided

(i) $m[n]^2 m[n-1] \leq \Theta M_n^{\alpha}$ for some $\alpha \in [0,l-4-\beta)$ independent of n,

(ii) $\eta(0) \leq \Xi$.

Furthermore $\eta(n) \leq \mathbf{C}^* \eta(0) M_n^{4+\beta-l}$, for a constant $\mathbf{C}^*$.

Recall that $\eta_l(n)$ denotes the sizes of the l-th derivatives of ν and τ after n renormalisations. The proof of the proposition splits into three parts:

(A) Under the assumption that $\eta_i(n) = \varepsilon_{i,0}(n) = O(\eta_l(n))$ for $1 \leq i < l$ uniformly in n we show in lemma 8 that $\varepsilon_i[u]$ are to be estimated by a polynomial in η and u, in a way that suggests that $\varepsilon_{i,u}(n) \sim u\eta_l(n)$ for $u \geq 0$. At this point we use the decomposition of the renormalisation operator introduced in section 2. In fact, lemma 8 deals only with the iterations of $N^{\wedge}$ and the estimates of $\varepsilon_{i,u}(n)$ are deduced inductively on i.

(B) Recall $N_m = N^* \circ N^{\wedge m}$. In lemma 8 we did the first part, i.e. $N^{\wedge m}$. Lemma 10 completes the renormalisation by applying N^*, which interchanges ν and τ and normlises the new pair using a coordinate transform Λ. The main part of Λ, an affine streching, makes N_m a contraction in the neigbourhood of the simple line. The non linear part of Λ will be introduced in lemma 10, and is designed to bring ν close to the affine shear ν_{∞}.

(C) In lemma 11 it will be demonstrated that the new pair of renormalised maps is well defined on $(\mathbf{C},\mathbf{f})$ and commutes in $\mathbf{f}$.

(D) Finally, to complete the proof of the proposition, we verify the hypothesis made in lemma 8. In particular it remains to prove that the new τ is close to an affine shear. To this end we use proposition 5 which chaines τ to ν provided they commute in some region, here $\mathfrak{F}$.

We fix n and drop the index n in the following where there is no likelihood of confusion. We use furthermore the notation $m = m[n]$ and $c = c[n]$. (The numbers $\mathfrak{r}_p$ will be the same as in proposition 5. In lemma 11 we shall return to this point.)

Lemma 8: In the definition of η_p let the unit internal $[0,1]$ be the range I of w. Suppose

(i) there are numbers $\mathfrak{r}_p$ so that $\eta_p \leq \mathfrak{r}_p \eta$, for $p = 2,\ldots,l-1$;

(ii) $(m+1)^2 c\eta_2 \leq \Theta^*$, where $\Theta^* = (54 \times 64)^{-1}$;

(iii) $\|\partial^p(\tau - \tau_{\infty,c})\|_{\mathbf{f}} \leq \eta_p$;

(iv) $\|\partial^p(\nu - \nu_{\infty})\|_{\mathbf{f}} \leq c^{-1}\eta_p$.

Then there are polynomials $\mathbf{Q}_p(x)$, $p = 2,\ldots,l$, with $\mathbf{Q}_p(0) = 0$ and positive coefficients such that

(4-1) $$\varepsilon_{p,u} \leq (u + c^{-1})\big(1 + \mathbf{Q}_p((u+1)^2 c\eta)\big)\eta_p$$

for $u = 0,\ldots,m$.

Remark: It the proof of proposition 7 it will turn out that $(m+1)^2c\eta$ (where $c \sim m[n-1]$) is a decreasing exponentially fast for $n \to \infty$, due to the fact that α has to be strictly less than $1-4-\beta$. Hence the factor $1 + \mathbf{C}_p((m+1)^2c\eta)$ will converge to 1. In particular, by shrinking Θ^* the infinite product of them over $n \in \mathbf{N}$ can be made as close to 1 as we please. The point of lemma 8 is therefore, that $\varepsilon_{p,u} \sim u\eta_p$.

Proof. We shall prove the lemma by induction on p. Beginning with second order derivatives, $p = 2$, we obtain for the derivatives of $\nu[u+1]$ in terms of $\nu[u]$ and $\tau[u]$:

$$D_w{}^2\nu[u+1](x,x') = D_w{}^2(\nu[u]\bullet\tau)(x,x')$$

$$= \nu_{20}[u] + w^2\tau_{02} + 2D_wx''[\nu_{11}[u] + w\tau_{11}] + (D_wx'')^2[\nu_{02}[u] + \tau_{20}]$$

where $u = 1,\ldots,m-1$, and where the ν's are evaluated at (x,x'') and the τ's at (x'',x'). (Because of the stationarity condition $\nu_{01}[u] + \tau_{10} = 0$ we do not get a factor involving $D_w{}^2x''$). Reordering gives

$$D_w{}^2\nu[u+1](x,x') = [\partial_1 + (D_wx'')\partial_2]^2\nu[u](x,x'') + (D_wx'')^2[\partial_2 + w(D_wx'')^{-1}\partial_3]^2\tau(x'',x'), \tag{4-2}$$

where ∂_1, ∂_2, ∂_3 denote partial derivatives with respect to x, x'', x'. When evaluating the square of $\partial_1 + (D_wx'')\partial_2$, the factor D_wx'' is regarded as a constant. Note that $N^\wedge$ does not touch τ at all. We will often drop the arguments of ν and τ. If c is chosen so that $\partial_1\partial_2(\tau-\tau_{\infty,c})|_{(0,0)} = 0$, then we say $\tau_{\infty,c}$ approximates τ best. Naturally, there are other ways to fix c which are no less suitable to prove the contraction of the renormalisation operator in the vincinity of the simple line. The derivatives $\{\nu_{20}, \nu_{11}, \nu_{02}\}$ can be expressed in terms of $\{D_w{}^2\nu\colon\ w \in [0,1]\}$, explicitly one finds (recall that $D_0 = \partial_1$)

$$\begin{aligned}
\partial_1{}^2 &= D_0{}^2,\\
\partial_1\partial_2 &= [4D_{1/2}{}^2 - D_1{}^2 - 3D_0{}^2]/2,\\
\partial_2{}^2 &= 2[D_1{}^2 - D_{1/2}{}^2 + D_0{}^2],
\end{aligned}$$

and reads off the inequality

$$\|\partial^2\nu[u]\| \le 8\|D_w{}^2\nu[u]\|, \tag{4-3}$$

for $w \in [0,1]$. Equipped with that inequality we do a closer look at $(D_w x^*)(x,x')$. From the stationarity condition $\nu_{01}[u](x,x^*) + \tau_{10}(x^*,x') = 0$ one derives with the implicit function theorem:

$$D_w x^* = -(\nu_{11}[u] + w\tau_{11})/(\nu_{02}[u] + \tau_{20})$$

$$= [c/(uc+1) + w + 0^*(8c\varepsilon_{2,u})]\times[c/(uc+1) + 1 + 0^*(8c\varepsilon_{2,u})]^{-1}$$

$$= 1 + (w-1)[(u+c^{-1})^{-1} + 1]^{-1} + 0^*(48c\varepsilon_{2,u})$$

since $c\varepsilon_{2,u} \leq 1/48$ by assumption (ii) (the denominator is expanded into a geometric series). The symbol 0^* has the following meaning: $f(x) = 0^*(x)$ if $|f(x)/x| \leq 1$ as $x \to 0$. The assumption

$$(w-1)[(u+c^{-1})^{-1} + 1]^{-1} \leq 1, \tag{4-4}$$

which will be made here shall be justified later on. We return to (4-2) and confine ourselves for the moment to pure D_w-derivatives. We forget for the time being about the change of the range or w when passing from u to $u+1$. This question will be discussed shortly. Hence

$$\varepsilon_{2,u+1} \leq \varepsilon_{2,u} + [2 + 48c\varepsilon_{2,u}]^2\eta_2 \leq \varepsilon_{2,u} + 9\eta_2, \tag{4-5}$$

if $48c\varepsilon_{2,u} \leq 1$. Similar to (4-3) where $\partial^2\nu$ was estimated by $D_w{}^2\nu$ for $w \in [0,1]$, the mixed ∂_2 and D_w derivatives can be estimated. For $u = 1$ we find $\varepsilon_{2,1} \leq c^{-1}\eta_2 + 9\eta_2$ since hypothesis (iv): $\|\partial^2(\nu-\nu_\infty)\| \leq c^{-1}\eta_2$. Summing over u we obtain:

$$\varepsilon_{2,u} \leq 9(u+c^{-1})\eta_2. \tag{4-6}$$

Let $[0,w[u]]$ be an interval contained in the range of w at the u-th renormalisation step. Especially in view of the restriction (4-4) we have to make shure that the domain of w in the (u+1)-st renormalisation step is included in the range of w in the u-th renormalisation step. More precisely, we shall show that

$$0 \leq D_{w[u+1]}x^* \leq w[u]$$

is satisfied. In fact this condition holds true for values w for which $0 \leq D_w x^* \leq w$ generally. This last inequality is clearly satisfied if $D_w x^* = [(u+c^{-1})^{-1} + w]\times[(u+c^{-1})^{-1} + 1]^{-1} + 0^*(48c\varepsilon_{2,u})$ lies in $[0,w]$, $\forall\ u = 0,\ldots,m$. This for example is true if

$$w \geq \max(1,\, 1 + 48\varepsilon_{2,u}(uc+1)),$$

and using (4-6) this follows from

$$w \geq 1 + 8^2 54(u+1)^2 c\eta_2,$$

which itself is satisfied by hypothesis (ii). It is not too hard to see, that the last inequality also implies $D_w x^* \geq 0 \;\; \forall \;\; u = 0,\dots,m$. We have therefore shown that the unit interval is contained in $[0,w[m]]$ provided $[0,w[0]] \subset [0,\, 1 + 54\times 64(m+1)^2 c\eta_2]$. In the sequel we shall assume $w[0] = 1 + 54\times 64(m+1)^2 c\eta_2$. In the hypothesis of the lemma η_p was supposed having been defined with w running over the unit interval. By the argument elaborated here, the unit interval $[0,1]$ has to be blown up to the size of $[0,w[0]]$. Naturally, the η_p then have to be replaced by some $\tilde{\eta}_p$, this time w running over $[0,w[0]]$. We return to this point at the end of the proof. In the following however we ignore this fact and assume η_p was defined by using $[0,w[0]]$ as range of w.

We have proven the lemma for $p = 2$. The general case is done via induction. Let us now assume formula (4-1) holds true for all derivatives up to $p-1$, where p is bigger than 2. We shall do the induction step from p to $p+1$. To express the p-th order derivatives of $\nu[u+1]$ in terms of $\nu[u]$ and $\tau[u]$ we need the following formula (where [illegible](s,t) stands for the binomial coefficient $s!/((s-t)!t!)$):

$$D_w^p \nu[u+1] = D_w^p(\nu[u] \circledast \tau)$$

(4-7)

$$= \sum_{0\leq q\leq p} \sum_{0\leq s\leq p-q} \partial_2^s \sum_{0\leq t\leq q} (\partial_1 + w\partial_3)^{q-t} \partial_2^t (\nu[u]\circledast\tau)(D_w x^*)^t \text{[illegible]}(s,t)\text{[illegible]}_p^{q,s}[u],$$

where the coefficients [illegible]$_p^{q,s}[u]$ are sums of monomials of the form

$$\prod_{1\leq r\leq s} D_w^{i[r]} x^*,$$

with $i[r] \geq 2$ and $\sum_{1\leq r\leq s} i[r] = p-q$. Identity (4-7) is proven by induction on p as follows. As we have seen in (4-2) the [illegible]s are trivial for $p = 2$. Assume (4-7) holds true for $p-1$ and lower orders, then there are three different ways for D_w to act:

(a) by increasing $s \to s+1$ and sending $t \to t-1$. The summands in the corresponding [illegible] then get one more factor;

(b) as partial derivatives directly on $(\nu[u]\circledast\tau)$: $t \to t+1$ and the [illegible] remain the same;

(c) it acts on the [illegible], increasing there the number of derivatives $\sum_r i[r]$ by one.

Note:

(α) stationarity makes the term $(t,s) = (0,1)$ vanish,

(β) the summand for $(t,s) = (1,1)$ is zero because $D_w x'' = -[\nu_{11}[u] + w\tau_{11}]/[\nu_{02}[u] + \tau_{20}]$, and

(γ) the only term with trivial $\mathfrak{P}$ is $(t,s) = (p,0)$.

In all other cases $\mathfrak{P}$ has at least one term of at least third derivative and itself is the coefficent of an at least third order derivative of $(\nu[u]\circledast\tau)$. Since third order derivatives of ν_∞ and $\tau_{\infty,c}$ are zero these summands are made up of products that have at least two factors each of them at least a third order derivative of $\nu[u]-\nu_\infty$ or $\tau[u]-\tau_{\infty,c}$ and are therefore majorized by a polynomial in η^2.

In formula (4-7) contraction of the sum over t gives

$$D_w{}^p\nu[u+1] = \sum_{0\le q\le p} \sum_{0\le s\le p-q} \{[\partial_1 + D_w x''\partial_2]^q \nu_{0,s}[u]$$

(4-8)
$$+ (D_w x'')^q [\partial_2 + (D_w x'')^{-1} w\partial_3]^q \tau_{s,0}\} \mathfrak{P}_p{}^{q,s}[u].$$

Unlike to the case $p = 2$ there is one more complication. On the right hand side of (4-2) appeared no mixed derivative $\partial_2 D_w$, that was because of the stationarity condition. For $p > 2$, additionally to $\{D_w{}^p(\nu[u] - \nu_\infty[u]): w \in [0,1],\ 2 \le p \le l\}$ we need to verify that for the derivatives $\{\partial_{x'}{}^q D_w{}^p(\nu[u] - \nu_\infty[u]): w \in [0,1],\ 2 \le p+q \le l\}$ the same estimates hold true as for pure D_w-derivatives. Consider $D_w{}^{p'}\nu[u+1]$, for some $p' < p$, and differentiate with respect to x', then there are two cases:

(i) $\partial_{x'}$ acts on τ, then the contribution is of size $O(\eta)$, since τ remains the same while u grows;

(ii) $(\partial_{x'}x'')\partial_{x''}$ acts of $\nu[u]$ and τ. We are finished when we can show that $|\partial_{x'}x''|$ is bounded by 1. Indeed

$$\partial_{x'}x'' = -\tau_{11}/(\nu_{02}[u] + \tau_{20}) = [(u+c^{-1})^{-1} + 1]^{-1} + O^*(48c\varepsilon_{2,u}),$$

which by hypothesis (ii) is less than one.

Suppose we were working with ∂_1, ∂_2 instead of D_w, ∂_2. Using the chain rule to evaluate the derivatives of $\nu[u]\circledast\tau$, each time we increase u to $u+1$, $\varepsilon_{*,u}$ has to be multiplied with a factor slightly bigger than one. If we replace ∂_1 by D_w we have instead to shrink the interval for w each time a bit. This procedure turns out to be technically more accessible.

From (4-7) we now pass on to estimate the ε's:

$$\varepsilon_{p,u+1} \le \varepsilon_{p,u} + [1 + (w-1)[(u+c^{-1})^{-1} + 1]^{-1} + 48c\varepsilon_{2,u}]^p\eta_p + \mathfrak{R}_{p,u}.$$

The remainder $\mathfrak{R}_{p,u}$ comes from those terms in (4-7) that involve non trivial polynomials $\mathfrak{P}$. As already noted they are dominated by sums of products with at least two $\varepsilon_{i,u}$, $3 \le i < p$. From (4-4) it follows that the second term in the brackets is less than one, and to estimate the entire coefficient of η_p we use the expression of $\varepsilon_{2,u}$ which we derived in the first part of the proof. Hence

$$\varepsilon_{p,u+1} \le \varepsilon_{p,u} + [1 + 48\times 9(u+c^{-1})c\eta_2]^p \eta_p + \mathfrak{R}_{p,u}$$

$$\le \varepsilon_{p,u} + [1 + 48\times 9e^{2p}(u+c^{-1})c\eta_2]\eta_p + \mathfrak{R}_{p,u},$$

where, in the second step, we got rid of the exponent p. We sum up over u:

(4-9) $$\varepsilon_{p,u} \le [(u+c^{-1}) + \eta_2 \textstyle\sum_{0\le j<u} 48\times 9e^{2p}(j+1)c]\eta_p + \sum_{0\le j<u} \mathfrak{R}_{p,j}.$$

Before all we finish off with the remainders. By induction hypothesis the $\varepsilon_{k,j}$ are for $k < p$ majorized by polynomials of the form

$$(j+1)a_1\eta + \textstyle\sum_{r\ge 2} (j+1)^{2r-1} a_r (c\eta)^r,$$

for some coefficients $\{a_r \ge 0:\ r \ge 1\}$. Products which contain at least two factors are dominated by something of the form

$$\textstyle\sum_{r\ge 2} (j+1)^{2r-2} a'_r (c\eta)^r,$$

with some new $\{a'_r \ge 0:\ r \ge 2\}$. Summing up over $j < u$, increases the exponent of $(j+1)$ by one (and multiplies with a factor $(j+1)^{-1} \le 1$). For the remainder in (4-9) we find

$$\textstyle\sum_{0\le j<u} \mathfrak{R}_{p,j} \le \sum_{r\ge 2} (j+1)^{2r-1} a'_r (c\eta)^r.$$

We add to the remaining terms of (4-9) (for simplicity substitute $(u+1)$ by $(u+1)^2$) and obtain a polynomial $\mathfrak{Q}'_p$ which has the properties as claimed for $\mathfrak{Q}_p$.

As earlier remarked, for initialising the iteration of $N^\wedge$ we have to consider the bounds on the D_w-derivatives where w runs over the interval $[0,\ 1 + 54\times 64(m+1)^2 c\eta_2]$. Furthermore, it is not too hard to see that there exists a (positive) polynomial $\mathfrak{P}$ satisfying

$$\sup_{i+j=p,w\in[0,w[0]]} (\|\partial_1{}^i D_w{}^j(\nu-\nu_\infty)\|,\ \|\partial_1{}^i D_w{}^j(\tau-\tau_{\infty,c})\|)$$

$$\le (1 + \mathfrak{P}(w[0]-1)) \sup_{i+j=p,w\in[0,1]} (\|\partial_1{}^i D_w{}^j(\nu-\nu_\infty)\|,\ \|\partial_1{}^i D_w{}^j(\tau-\tau_{\infty,c})\|).$$

We shall abstain from proving this formula. We need this inequality to get an estimate on the initial $\varepsilon_{p,0} = \eta_p$, which we have to replace by $\tilde{\eta}_p = (1 + \mathfrak{B}(54\times 64(m+1)^2 c\eta_2))\eta_p$ for $p = 2,\dots,l$. Without the dependency on m we would end up with unreasonably bad bounds in the case of small m. We go with this expression into $\mathcal{P}'_p$, and obtain the claimed polynomial $\mathcal{P}_p$. Obviously, $\mathcal{P}_p$ has the same properties as $\mathcal{P}'_p$. Positiveness of the coefficients and $\mathcal{P}_p(0) = 0$ follows from the construction. This completes the proof of the lemma. □

Before introducing the coordinate transformation Λ, we shall see how a map U' which is close to U_∞, can be 'flattened' using a non linear coordinate transform.

Lemma 9: Let $\rho_0 > 0$ be a constant and (U',T') be near enough to the simple line, and let (ν',τ') be their generating functions defined on $(\mathcal{C},\mathcal{F})$. Suppose there are $\rho \in (0,\rho_0]$ and $\Gamma_x \geq 1$ so that

$$\|\partial^p(\nu'-\nu_\infty)\|_{\mathcal{C}} \leq \rho\Gamma_x^{-p}$$

for $1 \leq p \leq l$. Then there exists a coordinate change h and constants $\mathfrak{s}, \mathfrak{t} > 0$ independent of (ν,τ), but depending on ρ_0, so that

(i) (a) $\|\partial^p(h\circ\nu'\circ h^{-1}-\nu_\infty)\|_{\tilde{\mathcal{C}}} \leq \rho\mathfrak{s}^{l-p}$ for $p = 1,\dots,l-1$;

(b) $\partial^l(h\circ\nu'\circ h^{-1} - \nu') = 0$,

where $\tilde{\mathcal{C}}$ is a region in the x,x'-plane slightly smaller than $\mathcal{C}$ and converges to $\mathcal{C}$ for $\rho_0 \to 0$.

(ii) For $\tau'(x,x')$ we get

(a) $\|\partial^p(h\circ\tau'\circ h^{-1} - \tau')\|_{\mathcal{F}} \leq \mathfrak{t}\rho\Gamma_x^{-p}$ for $p = 1,\dots,l-1$;

(b) $\partial^l(h\circ\tau'\circ h^{-1} - \tau') = 0$.

Proof. The coordinate transform h depends entirely on ν' and so we will not talk about τ' until the end of the proof.

Coordinate transformations h that transform area preserving maps again into area preserving maps must have constant $\det Dh$, where Dh is the Jacobian of h. This means that up to some affine stretching, h has to be area preserving itself. If h has the twist property, then it can be represented by a generating function. We shall go this way, combining a non linear coordinate transform with a shift by some factor close to $1/2$ to obtain a coordinate transform h which has the twist property and therefore can be represented using a generating function. The shearing factor $1/2$ is chosen so that the compositions of h and h^{-1} with U and T are again shearings and possess therefore well defined generating functions.

Set h_∞: $(x,y) \to (x',y') = (x+2^{-1}y,\, y)$ which has a generating function $\mu_\infty(x,x') = (x-x')^2$. Let h_0 be a coordinate transform that is generated by $\mu(x,x')$, then h_0^{-1} is generated by $\mu_0^{-1}(x,x') = -\mu_0(x',x)$. In fact we shall determine the generating function of h_0, rather than giving h_0 in x,y-coordinates. Put $U^{\sim} = h_0 \circ U' \circ h_0^{-1}$ and introduce variables according to the diagram:

$$(x,y) \overset{h_0}{\twoheadrightarrow} (X,Y) \overset{U'}{\twoheadrightarrow} (X',Y') \overset{h_0^{-1}}{\twoheadrightarrow} (x',y'),$$

where $U^{\sim}$: $(x,y) \to (x',y')$. In particular we obtain

$$\nu^{\sim}(x,x') = \mu_0(x,X) + \nu'(X,X') - \mu_0(x',X')$$

for intermediate coordinates X,X'. To determine $\mu_0(x,x')$ we shall replace X,X' by X^*,X'^* by which we denote the intermediate coordinates in the purely affine composition

$$\nu_\infty^{\sim}(x,x') = \mu_\infty(x,X^*) + \nu_\infty(X^*,X'^*) - \mu_\infty(x',X'^*).$$

Considering x',X'^* as variables we obtain $X^* = X^*(x,x')$, $X'^* = X'^*(x,x')$ and since $X'^*-X^* \le -4^{-1}$ it follows in particular $x'-x \le -8^{-1}$ if η is small enough. Take the Taylor expansion of $(\nu'-\nu_\infty)(X^*,X'^*)$, cut it off at terms of order l and call $\upsilon(X^*,X'^*)$ the resulting polynomial of order $l-1$. Define

$$\mu_0(x',X'^*) = \mu_\infty(x',X'^*),$$
$$\mu_0(x,X^*) = \mu_\infty(x,X^*) - \upsilon(X^*,X'^*) \quad \text{for} \quad x(x',X'^*), X_0(x',X'^*),$$

where $x'+X'^* = -2$. Since X'^*-X^*, $x'-x \le -1/8$ there is enough room for smoothing $\mu_0(x,X^*)$ so that

(i) is C^l, and

(ii) $\mu_0(x,X^*) = \mu_\infty(x,X^*) - \upsilon(X_0,X'^*) - \mu_0(x',X'^*)$,

where $(x,X^*) = (x(x',X'^*), X^*(x',X'^*))$, so that $(X^*,X'^*) \in \mathbf{C}$. Note that (ii) defines $\mu_0(x,X^*)$ uniquely, provided the smoothing from $x'+X'^* = -2$ to $\{(x(x',X'^*), X^*(x',X'^*))\colon x'+X'^* = -2\}$ is done. Unfortunatly the intermediate points X,X' of

$$\nu^{\sim}(x,x') = \mu_0(x,X) + \nu(X,X') - \mu_0(x',X')$$

in gerneral will be different from X^*,X'^*, but since

$$\|\partial^p(\mu_0-\mu_\infty)(x,x')\| \le \|\partial^p\upsilon(x,x')\| \le \rho\Gamma_x^{-p},$$

we are already pretty close, namely

$$|X-X^*|, |X'-X'^*| \le *\|\partial\upsilon(x,x')\| \le *\rho\Gamma_x^{-1}.$$

for some $\varkappa > 0$ independent of ν'. This follows from the stationarity conditions for X, X' and which can be written as the stationarity codition of the affine maps plus error terms of size $O(\rho\Gamma_x^{-1})$. From this we get by integration

$$\begin{aligned}
\mu_{0;i,j}(x,X) &= \mu_{0;i,j}(x,X^*) + {}_{X^*}\int^{X} \mu_{0;i,j+1}(x,\zeta)d\zeta, \\
\mu_{0;i,j}(x',X') &= \mu_{0;i,j}(x',X^*) + {}_{X^*}\int^{X'} \mu_{0;i,j+1}(x',\zeta)d\zeta, \\
\nu'_{i,j}(X,X') &= \nu_{\infty;i,j}(X^*,X^*) + {}_{X^*}\int^{X} \upsilon_{i,j+1}(\zeta,X^*)d\zeta + {}_{X^*}\int^{X'} \upsilon_{i,j+1}(X,\zeta)d\zeta,
\end{aligned}$$

and in particular since the paths of integration are of length $\leq \varkappa\rho\Gamma_x^{-1}$ we find

$$\nu^{\sim}_{i,j}(x,x') = \mu_{0;i,j}(x,X^*) + \nu_{\infty;i,j}(X^*,X^*) - \mu_{0;i,j}(x',X^*) + O^*(4\varkappa\rho^2\Gamma_x^{-i-j-1}),$$

where the symbol O^* has the same meaning as in lemma 8, i.e. the remainder term is estimated by exactly $4\varkappa\rho^2\Gamma_x^{-i-j-1}$. Therefore

$$\|\nu^{\sim}_{i,j}(x,x')\| \leq 4\varkappa\rho^2\Gamma_x^{-i-j-1},$$

for $i+j > 2$. The norm here is taken over a suitable region in the x,x'-plane. The corresponding region in the x,y-plane is $\mathfrak{C}^*$ (introduced in section 2) shifted by h_0. This improves the deviation of the derivatives of $\nu'-\nu_\infty$ from $\rho\Gamma_x^{-i-j}$ to $4\varkappa\rho^2\Gamma_x^{-i-j-1}$ for $\nu^{\sim}-\nu_\infty$ by approximately a factor Γ_x^{-1}. The final coordinate transform h is found by iterating this process as follows. Replace ν by $\nu^{\sim}$ and we proceed to construct a second transform h_1 in exactly the same way as we determined h_0. However there will be one difference, unlike h_0 which is close to an affine shear with factor $-1/2$, h_1 will be close to the affine shear $h_{-\infty} = (x-y/2, y)$. Hence $\mu_\infty = (x-x')^2$ has to be substituted by $\mu_{-\infty} = -(x-x')^2$. So the composition $h_1 \circ h_0$ is close to the identity. Similar to the estimations elaborated above for h_0, we obtain (since $\rho^2 \leq \rho$)

$$\|(h_1 \circ h_0 \circ \nu \circ h_0^{-1} \circ h_1^{-1})_{i,j}(x,x')\| \leq (4\varkappa\rho)^2\rho\Gamma_x^{-i-j-2}.$$

Here one takes the norm over some region $\mathfrak{C}_1$, which approximately coincides with $\mathfrak{C}$. We shall not go into details, but it is clear that $\mathfrak{C}_1$ containes at least the quatrilateral that one obtaines from $\mathfrak{C}$ by cutting off on all sides a strip of size $O(\rho\Gamma_x^{-1})$. Going on we find h_2 close to h_∞, and in general h_k, $k = 0,\dots,l$, which is close to the shear $(x,y) \to (x',y') = (x+(-1)^k y/2, y)$. If l happens to be odd the we set $h = h_l \circ \dots \circ h_1 \circ h_0$ and in the case that l is an even number we set $h = h_{l+1} \circ h_l \circ \dots \circ h_0$, where $h_{l+1} = (x',y') = (x-y/2, y)$ an affine shear designed to make h itself close to the identity. Clearly, under the transformation with h_{l+1} one does not affect the non linear parts of ν. Furthermore from the remark made in the last paragraph it is obvious that $h\circ\nu'\circ h^{-1}$ is defined in some region slightly smaller than $\mathcal{C}$, and can be made arbitrary close to $\mathcal{C}$ itself by shrinking the constant ρ_0.

The statement (i,b) follows from the fact that the coordinate transforms h_k are polynomials of order less that l-1.

To finish off let τ' be any generating function. Put $\tau^\sim = h\circ\tau'\circ h^{-1}$ and we obtain in the same way as above

$$|\partial^p(h_0\circ\tau'\circ h_0^{-1} - \tau)|_{\mathcal{C}^\wedge} \le |\partial^p(\mu_0-\mu_\infty)|_{\mathcal{C}^\wedge} \le \rho\Gamma_x^{-p},$$

for $p = 2,\dots,l-1$, where $\mathcal{C}^\wedge$ is some region in the x,x'-plane which is determined by ν; and $\partial^l(h_0\circ\tau'\circ h_0 - \tau') = 0$ since $\partial^l(\mu_0-\mu_\infty) = 0$. We shall be moe precise about $\mathcal{C}^\wedge$. First note that saying ν' is defined in $\mathcal{C}$ is the same as to say U' is defined in $\mathcal{C}^*$, where $\mathcal{C}^*$ is in the (x,y)-plane up to an error of size $O(\rho\Gamma_x^{-1})$ the region

$$\{(x,y) \in \mathbb{R}^2:\ -3 \le y \le 3/4,\ |x - (1-y)/2| \le 2\}$$

Let T': $(x,y) \to (x',y')$, (where $(y,y') = (-\tau'_{10}(x,x'), \tau'_{01}(x,x'))$) be defined in $\mathcal{F}^* \subset \mathbb{R}^2$. Since we are close to the simple line as we wish, we may assume that $\mathcal{F}^* \subset \mathcal{C}^*$ and can therefore replace $\mathcal{C}^\wedge$ by $\mathcal{F}$. Iterating yields

(4-10) $$|\partial^p(h\circ\tau'\circ h^{-1} - \tau')|_{\mathcal{F}} \le \rho\Gamma_x^{-i-j}\sum_{0\le k<l}(4\varepsilon\rho)^k\Gamma_x^{-k} \le (1+\varepsilon)\rho\Gamma_x^{-i-j},$$

where $\varepsilon = \sum_{1\le k<l}(4\varepsilon\rho_0)^k$ is a constant independent of (ν',τ'), and can be made arbitrary small if ρ_0 is small enough. It is clear that $\partial^l(h\circ\tau'\circ h^{-1} - \tau') = 0$ and this concludes the proof. □

In lemma 8 we examined the effect that $N^{\wedge m}$ has on (ν,τ). To complete the renormalisation N_m we have to apply N^* which maps $(\tau,\nu[m]) \to (\Lambda\circ\nu[m]\circ\Lambda^{-1}, \Lambda\circ\tau\circ\Lambda^{-1})$, where the coordinate change Λ is designed to bring $(\tau,\nu[m])$ into normal form. In the following lemma we shall determine Λ. This is done in two steps. The first consists in giving the affine part Λ', which up to shifting in y-direction streches in x and y direction. In the second step ν gets 'flattened' whereby we need lemma 9.

Lemma 10: Let $\Delta = \Delta_n$ be as in proposition 7 and set $\Delta_- = \Delta_{n-1}$, $m = m[n]$, $c = c[n]$, $\eta = \eta(n)$ and $\eta' = \eta(n+1)$. Then there exists a coordinate transformation Λ (= Λ_n) close to affine and a constant ε^* (independent of (ν,τ), so that

(i) $\eta' \le (1 + \varepsilon^* m^2 c\eta)\cdot\Delta^{l-4-\beta}\eta$;

(ii) $|\partial^p(\tau^* - \tau_{\infty,c})|_{\gamma} = O(\Delta)$ for $p = 2,\dots,l-1$;
(iii) $|\partial^p(\nu^* - \nu_{\infty})|_{\epsilon} = O(\eta')$ for $p = 2,\dots,l-1$;
(iv) $|\partial^l\tau^*|_{\gamma} = O(\eta')$;
(v) $|\partial^l\nu^*|_{\epsilon} = O(\Delta\eta')$;
where $\tau^* = \Lambda\circ\nu[m]\circ\Lambda^{-1}$, $\nu^* = \Lambda\circ\tau\circ\Lambda^{-1}$ are the transformed generating functions of U_n, T_n.

Proof. The coordinate change Λ falls apart into two parts, $\Lambda = h\circ\Lambda'$, where Λ' is affine. The affine part normalizes $(\nu[m],\tau)$. Its Jacobian looks in x,y - coordinates like

$$D\Lambda' = \begin{vmatrix} -\Gamma_x & 0 \\ 0 & -\Gamma_y \end{vmatrix},$$

for positive numbers Γ_x, Γ_y. We have in y-direction

(4-10-y) $$\Gamma_y = [(mc+1)^{-1} + O^*(\varepsilon_{2,m})]^{-1} = (m+c^{-1})c + O^*(48mc\varepsilon_{2,m}),$$

and in x-direction

(4-10-x) $$\Gamma_x = \Gamma_y(c^{-1} + O^*(\eta_2)) = m + c^{-1} + O^*(64mc\varepsilon_{2,m}).$$

Set $(U',T') = (\Lambda'\circ T\circ\Lambda'^{-1}, \Lambda'\circ U[m]\circ\Lambda'^{-1})$; we summarize:
(i) U is near to a shearing by factor one,
(ii) T is near to a shearing by $c[n]$, and
(iii) $U[m]$ is up to an error of size $\varepsilon_{2,m}$ close to a shearing by $mc+1$.
Hence
(α) $U' = \Lambda'\circ T\circ\Lambda'^{-1}$ is a shearing by one (up to an error $O^*(\eta'_2)$, where $\eta'_2 = \eta'_2(n+1)$),
(β) $T' = \Lambda'\circ U[m]\circ\Lambda'^{-1}$ shears by

$$c' = \Gamma_x\Gamma_y^{-1}[mc + 1 + O^*(\varepsilon_{2,m}) = [m + c^{-1} + O^*(64mc\varepsilon_{2,m})](1 + O^*(36\times 8\varepsilon_{2,m})).$$

Thus, c' can be approximated by Δ^{-1}, more precisely

(4-11) $$\log c'\Delta = O^*(45mc\varepsilon_{2,m}) \le 1/2$$

and Λ' transforms generating functions like:

$$\tau(x,x') \to \tau'(x,x') = (\Lambda'\circ\tau\circ\Lambda'^{-1})(x,x') = \Gamma_x\Gamma_y\tau(-\Gamma_x^{-1}x, -\Gamma_x^{-1}x') + \text{const.}(x+x'),$$

where the constant expresses the contribution from shifting in y-direction.

Similarly for $\nu'(x,x') = (\Lambda'\circ\nu\circ\Lambda'^{-1})$. The p-derivatives of the transformed functions are multiplied by a factor which is in modulus $\Gamma_y\Gamma_x{}^{1-p}$. We note that the ν-derivatives turn out to be smaller by a factor Γ_x^{-1} that the τ-drivatives. This is because τ remains untouched under the iteration of $N^{\wedge}$, i.e. $\tau[u] = \tau$ for all $u = 0,\ldots,m$. Hence

(4-12-τ) $\max_{w\in I, j+k=p} \|\partial_2{}^j D_w{}^k(\tau'-\tau_{\infty,c})\| \le \Gamma_y\Gamma_x{}^{1-p}\varepsilon_{p,m},$

(4-12-ν) $\max_{w\in I, j+k=p} \|\partial_2{}^j D_w{}^k(\nu'-\nu_{\infty})\| \le \Gamma_y\Gamma_x{}^{1-p}\eta_p,$

for $1 \le p < l$. We shall now determine the non linear part h of the coordinate transform Λ. In lemma 9 put $\rho = \Gamma_y\Gamma_x\eta_p$ and identify the Γ_x (conspiciously the same notation). Statement (ii) of lemma 10 follows immediatly. It remains to check the statement concerning τ. Here we use that h changes the Taylor polynomial of generating function to not higher than $(l-1)$st order. In particular, the l-th order derivatives of $\nu^* = h\circ\nu'\circ h^{-1}$, $\tau^* = h\circ\tau'\circ h^{-1}$ are the same as these of ν', τ'. More precisely (with the constant $\mathbf{t}$ from lemma 9)

(i) $\max_{w\in I, j+k=p} \|\partial_2{}^j D_w{}^k(\tau^*-\tau_{\infty,c})\| \le \Gamma_y\Gamma_x{}^{1-p}(\varepsilon_{p,m} + \mathbf{t}\eta_p)$ if $p < l$;

(ii) $\max_{w\in I, j+k=l} \|\partial_2{}^j D_w{}^k(\tau^*-\tau_{\infty,c})\| \le \Gamma_y\Gamma_x{}^{1-l}\varepsilon_{l,m}.$

For the sake of completeness we listed (i), it is necessary to persue only the second case. Put η'_l for $\eta_l(n+1)$ and estimate $\varepsilon_{l,m}$ with the assistance of (4-1) from lemma 8. We find

(4-13) $$\eta'_l \le (m+c^{-1})\Gamma_y\Gamma_x{}^{1-l}[1 + \mathbf{O}_l(m^2c\eta)]\eta_l$$

To evaluate the factor $(m+c^{-1})\Gamma_y\Gamma_x{}^{-3}$ we need (4-10-y), (4-10-y) and (4-11). Firstly

$$(m+c^{-1})\Gamma_x^{-1} \le (m+\Delta_-)\Delta + 64mc\varepsilon_{2,m} + 45c\varepsilon_{2,m} \le 1 + \Delta\Delta_- + 109mc\varepsilon_{2,m}.$$

Furthermore we use the property $\Gamma_x^{-1} \sim \Delta$, $\Gamma_y^{-1} \sim \Delta_-$; more precisely, we collect from (4-10-y) and (4-10-y):

$$\Gamma_y\Gamma_x{}^{-2} = \Gamma_y^{-1}[c^{-1} + O^*(\eta_2)]^{-2} = \left[(m+c^{-1})c^{-1} + O^*(3(m+c^{-1})\eta_2)\right]^{-1}$$

$$= c(m+c^{-1})^{-1} + O^*(3c\eta_2) = \Delta_-{}^{-1}\Delta + O^*(4c\eta_2).$$

From the last two estimates we gather

$$(m+\Delta_-)\Gamma_y\Gamma_x{}^{-3} = \Delta_-{}^{-1}\Delta(1+\Delta\Delta_-) + O^*(121mc\varepsilon_{2,m}),$$

and

$$\eta'_l \le c\Delta^{l-4}\Delta_-^{-1}\Delta(1+\Delta\Delta_-)\eta_l,$$

for a constant

$$c = c_n = 1 + (\mathfrak{a}_p(m^2c\eta) + 121mc\varepsilon_{2,m} + 64mc\varepsilon_{2,m})\Gamma_x^{-1}.$$

Since $\varepsilon_{2,m} \le 9(m+c^{-1})\eta_2$, the terms in brackets are all of size $m^2c\eta$ and henceforth there exists a constant $\mathfrak{c}^* > 0$ so that $\mathfrak{c} = \mathfrak{c}_i \le 1 + \mathfrak{c}^*m^2c\eta$.

To verify the statement (iv) we shall do a closer look at the factor $(1+\Delta\Delta_-)$. For that purpose introduce renormalisation indices n. If $m[n]$, $m[n-1]$ are large numbers, $\Delta = \Delta_i$ and $\Delta_- = \Delta_{i-1}$ will be close to zero. In the case where $m[n]$ increases fast enough for $n \to \infty$ (at least like n^θ for some $\theta > 1/2$), the product $\prod_{i<n}(1+\Delta_i\Delta_{i-1})$ converges to some finite value for $n \to \infty$. However this need not necessarily to be true, and in particular is not true if the rotation number ω we are dealing with is of constant type. A priori we can only say that $(1+\Delta_i\Delta_{i-1})$ is at most 2, since Δ_i may come as close to 1 as one may fear. This, however happens only when $m[n] = 1$ and $m[n-1]$ is some large number; then $\Delta_{i-1} \sim m[n-1]^{-1}$, $\Delta_i \sim 1$ and Δ_{i+1} is less than $1/2$ whatever $m[n+1]$ happens to be. This argument shows, that subject to some fluctuations, the products $\prod_{1\le i<n}(1+\Delta_i\Delta_{i-1})$ can be estimated by setting Δ_i the inverse of the golden mean $\mathfrak{g} = (1+\sqrt{5})/2$. Hence $(1+\Delta_i\Delta_{i-1}) \le (1+\mathfrak{g}^{-2})$ except in the case just described. If $\beta = \log(1+\mathfrak{g}^{-2})/\log\mathfrak{g}$ (which is less than 1) we have $(1+\mathfrak{g}^{-2})\mathfrak{g}^{-\beta} \le 1$. The gist of this is the statement (i):

$$\eta'_l \le \mathfrak{c}\Delta^{l-4-\beta}\eta_l.$$

To level out the fluctuations which are caused when in the continued fraction expansion of ω when $m[n]$ of too different size are too near, we should additionally to $\mathfrak{c}$ introduce one more constant $\mathfrak{b} = \mathfrak{b}_i$. As we pointed out in the last paragraph, the product $\prod_{0\le i<n}\mathfrak{b}_i$ remaines close to one for $n \in \mathbb{N}$; in particular is uniformly bounded (e.g. by 2). For clarity we shall ignore this fact as we already did in the statement of the lemma.

In the last paragraph statement (iv) $\|\partial^l\tau^*\| = O(\eta')$ was proven. Statement (v), $\|\partial^l\nu^*\| = O(\Delta\eta')$, follows immediately from (4-12-ν). Furthermore, the non linear coordinate transform h brought ν tclose to an affine shear, i.e. $\partial^p(\nu^*-\nu_\infty)_{i,j}$, therefore $\|\partial^p(\nu^*-\nu_\infty)_{i,j}\| = O(\|\partial^l\nu^*\|) = O(\Delta\eta')$, $p = 2,\dots,l-1$, and hence statement (iii).

It remains to verify satement (ii), that the derivatives $\partial^p(\tau^*-\tau_{\infty,c})$ are of size $O(\Delta)$ for $p = 2,\dots,l-1$. Recall (4-12-τ): $\|\partial^3(\tau'-\tau_{\infty,c})\| \le \mathfrak{p}\Gamma_y\Gamma_x^{-2}\varepsilon_{3,m}$, this in conjuction with $\varepsilon_{3,m} \le \mathfrak{q}m\eta$, where $\mathfrak{p}, \mathfrak{q}$ are constants independent of m and η. Without going into detail we use the property $\Gamma_x \sim m \sim \Delta^{-1}$, $\Gamma_y \sim c \sim \Delta_-^{-1}$. By lemma 8 we have $cm^2\eta = O(1)$. Combine these estimates in the order

$\varepsilon_{p,m} = O(\eta'_p c^{-1} m^{p-2})$, $\eta = O(m^{-1}\varepsilon_{p,m})$, $cm^2\eta = O(1)$ and one finds $\eta'_p m^{p-1} = O(1)$, which by far satisfies $\|\partial^p(\tau'-\tau_{\infty,c})\| = O(\Delta)$ for $p = 2,\ldots,l-1$. Hence the lemma.

□

Remark: Let $h = h_{l'}\circ\ldots\circ h_0$ (where l' is either l or $l+1$) be the nonlinear part of the coordinate transform Λ and denote by $\upsilon[i]$ the generating function of h_i, for $i = 0,\ldots,l'$. Then it follows from lemma 9 in conjunction with lemma 10

$$\|\partial^p\upsilon[i]\| \le \rho\Gamma_x{}^{-p} \le \Gamma_y\Gamma_x{}^{1-p}\eta_p.$$

For the non linear coordinate transform $h: (x,y) \to (x',y') = (x + F(x,y),\ y + G(x,y))$ in the x,y-plane we collect (one uses the implicit function theorem to compute the derivatives of F, G) the following inequalities

$$\|\partial^p F\|,\ \|\partial^p G\| \le (1+t^*)\rho\Gamma_x{}^{-p-1} \le (1+t^*)\Gamma_y\Gamma_x{}^{-p}\eta_{p+1}$$

$\forall$ $p = 1,\ldots,l-1$ for some constant $t^* > 0$, independent of (ν,τ) but depending on t. In particular $t^* \to 0$ for $t \to 0$, i.e. if ρ_0 goes to zero, where ρ_0 itself depends on Θ^*. Hence t^* can be made arbirtary small for Θ^* small enough.

Lemma 11: Suppose (ν,τ) is a normalised pair defined on $(\mathfrak{C},\mathfrak{F})$. Then the renormalised pair $N_m(\nu,\tau)$ is again well defined in $(\mathfrak{C},\mathfrak{F})$, commutes in $\mathfrak{F}^*$.

Proof. We consider the composition

$$\nu[u+1](x,x') = \nu[u](x,x'') + \tau(x'',x').$$

The stationarity condition for x, x'', x' is

$$-(x-x''-1)/(uc+1) + (x''-x')/c = O^*(\varepsilon_{2,u}+\eta_2).$$

Regrouping yields

$$\text{(4-14)}\qquad \begin{aligned} x-x' &= 1 - (x''-x')(u+1+c^{-1}) + O^*(10c(u+1)^2\eta_2),\\ x''-x' &= ((x-x'') - 1)(u+c^{-1})^{-1} + O^*(10c(u+1)\eta_2), \end{aligned}$$

where we used $\varepsilon_{2,u} \le 9(u+1)\eta_2$ which was proven in lemma 8. Let $[s_{u+1,-}, s_{u+1,+}]$ be the range of $x''-x'$. Clearly $[s_{0,-}, s_{0,+}] = [\max(-2\sqrt{t}, -3c/4), \min(2\sqrt{t}, 2c)]$, and from (4-14) we derive inductively

$$s_{u+1,+} = \min(2\sqrt{t}, 2c, (r_{u,+} - 1)(u-1+c^{-1})^{-1} - 10cu\eta_2),$$

$$s_{u+1,-} = \max(-2\sqrt{3/2},\ -3c/4,\ (r_{u,-} - 1)(u-1+c^{-1})^{-1} + 10cu\eta_2),$$

for $u = 1,\dots,m$, where $[r_{u,-}, r_{u,+}]$ is the range of $x-x''$ (with the notation introduced for $\nu[u+1] = \nu[u]\circledast\tau$), for which we obtain (also from (4-14))

$$r_{u,+} = 1 - s_{u,-}(u+1+c^{-1}) - 10c(u+1)^2\eta,$$
$$r_{u,-} = 1 - s_{u,+}(u+1+c^{-1}) + 10c(u+1)^2\eta_2,$$

where $[r_{0,-}, r_{0,+}] = [1/4, 3]$. For $u = 1$ we find

$$s_{1,+} = \min(2\sqrt{3/2},\ 2c,\ 2c - 10c\eta_2),$$
$$s_{u,-} = \max(-2\sqrt{3/2},\ -3c/4,\ -3c/4 + 10c\eta_2),$$

and

$$r_{1,+} = \min(1 + 3c/2 - 20c4\eta_2,\ 1 + 4\sqrt{3/2} - 10c4\eta_2)$$
$$r_{1,-} = \max(1 - 4c + 20c4\eta_2,\ 1 - 4\sqrt{3/2} - 10c4\eta_2).$$

By shrinking the constant Θ^* which we introduced in lemma 8 we can achieve in particular that (i) $10c(u+1)^2\eta_2$ is arbirary small, and (ii) $c \geq 1-\iota$, where $\iota > 0$ (independent of n) can be made arbitrary small. Thus we can assume that $r_{1,+} \geq 2$, and $r_{1,-} \leq -2$. It is not too hard to see that $r_{u,+}$ is monotone increasing, i.e. $r_{u+1,+} \geq r_{u,+}$, and similarly $r_{u+1,} \leq r_{u,}$ for $u = 1,\dots,m-1$. Hence $\nu[u]$ is well defined for $x-x' \in [-2,2]$, $u = 1,\dots,m$.

To equation (4-14) we add $2x' = (x''+x') - (x''-x')$ and obtain

$$\text{(4-14)} \qquad x+x' = 1 - (x''-x')(u+2+c^{-1}) + (x''+x') + O^*(10c(u+1)^2\eta_2),$$

Since u is at least 0, and $10c(u+1)^2\eta_2$ is arbitrary small, we see (without further elaborating) that the range of $x+x'$ is at least the range of $x''+x'$. Hence $\nu[u]$ is defined for $|x+x'| \leq 4\ \forall\ u = 1,\dots,m$.

Finally, let $\nu[m]$ be defined on $\{(x,x') \in \mathbb{R}^2\colon |x-x'| \leq 2,\ |x+x'| \leq 4\}$. If $|\Gamma_x| \geq \sqrt{3/2}$ then it is clear that $\Lambda' \circ \nu[m] \circ \Lambda'^{-1}$ is defined on $\mathcal{F} = \{(x,x') \in \mathbb{R}^2\colon |x-x'| \leq 2\sqrt{3/2},\ |x+x'| \leq 4\}$. (We neglect the transform h since its effect can be estimated in terms of $c(u+1)^2\eta_2$, can therefore be made arbitary small by way shrinking Θ^*.) Futhermore, it is clear that $\mathcal{C} \subset \Lambda'(\mathcal{F})$ and $\mathcal{F} \subseteq \Lambda'(\mathcal{C})$ (since $3 \leq 2\sqrt{3/2}\sqrt{3/2}$).

The condition $|\Gamma_x| \geq \sqrt{3/2}$ causes a problem in the case a large $m[n-1]$ is followed up by a small $m[n] = 1$. In that case we have $\Delta_n = (1+\Delta_{n-1})^{-1} \sim 1$, where $\Delta_{n-1} \sim m[n-1]^{-1}$ is nearly zero. We discussed this problem in lemma 10. The same argument shows that $\sqrt{3/2}$ is a good lower bound for $|\Gamma_x|$ except when $\Gamma_x \sim \Delta_n = (1+\Delta_{n-1})^{-1} \sim 1$. In that case we replace $\mathcal{F}$ by a somewhat smaller region and one renormalisation step later we return to the full size $\mathcal{F}$.

It is clear, that $\nu[u]$ and τ commute in $\{(x,x') \in \mathbb{R}^2\colon |x-x'| \leq 2, |x+x'| \leq 4\}$ $\forall$ $u = 1,\ldots,m$, and therefore, $\Lambda'\circ\nu[m]\circ\Lambda'^{-1}$ and $\Lambda'\circ\tau\circ\Lambda'^{-1}$ commute in $\mathfrak{f}$.

By making the constants Θ^* small enough we can achieve that the effect of the non linear coordinate transform h on the region $\mathfrak{C}$ is arbitrary small. We shall abstain from elaborating this fact, and note only that we are entiteled to replace Λ' by Λ without having to change the statements. Hence the lemma. $\square$

Proof of proposition 7. Iterating the statement (i) of lemma 10, we find

$$\eta_l(n) \leq \mathfrak{C}_n M_n^{4+\beta-l}\eta_l(0),$$

where $\mathfrak{C}_n = \prod_{0\leq i<n} \mathfrak{c}_i$. Let $\Theta^* \leq (54\times64)^{-1}$ be a constant so that $1+\mathfrak{c}^*\Theta^* \leq \mathfrak{z}^{\delta/4}$, say. Tighten the hypothesis (i) of lemma 8 in a way, so that $m[n]^2m[n-1]\eta_2 \leq \Theta^*$ is satisfied. Hence $\mathfrak{C}_n$ is less than $\mathfrak{z}^{n\delta/4}$ $\forall$ $n \in \mathbb{N}$. Set $\delta = l-4-\beta-\alpha$, then $\delta > 0$ since α is supposed to be strictly less than $l-4-\beta$. Then

$$m[n]^2c[n]\eta_2(n) \leq \mathfrak{c}_2 m[n]^2c[n]\eta(n) \leq \mathfrak{c}_2\Theta^*\mathfrak{C}_n M_n^{-\delta}\eta(0),$$

because $m[n]^2c[n] \leq \Theta M_n^{\alpha}$ by hypothesis (i) of the proposition. Since M_n increases at least exponentially, for instance $M_n \geq \mathfrak{z}^{n/2}/2$ (since we are allowed to assume $\Delta_i \leq \sqrt{\mathfrak{z}}^{-1}$ $\forall$ $i \in \mathbb{N}$, see lemma 10), the right hand side decreases exponentially fast like const.$\mathfrak{z}^{-n\delta/4}$ for $n \to \infty$. Hence, the product

$$\mathfrak{C}_n = \prod_{0\leq i<n} \mathfrak{c}_i \leq \exp \sum_{n\in\mathbb{N}} \mathfrak{c}^* m[n]^2c[n]\eta(n)$$

converges to some finite value $\mathfrak{C}^*$, $\mathfrak{C}_n \leq \mathfrak{C}^*$. (One can give an upper bound on $\mathfrak{C}^*$ that does not depend on (ν,τ) but on Θ^*.) We get the convergence result

$$\eta(n) \leq \mathfrak{C}^*\eta(0)M_n^{4+\beta-l}.$$

In particular, to meet the condition $m[n]^2c[n]\eta_2(n) \leq \Theta^*$ for $n = 0$, we set $\Xi = \Theta^*(\mathfrak{C}^* m[0]^2c[0])^{-1}$, depending on T_0 and ω. Set $\Theta = \Theta^*(\mathfrak{C}^*\mathfrak{c}_2\Xi)^{-1}$, and using hypothesis (i) of proposition 7, $m[n]^2m[n-1]\Theta^{-1}M_n^{-\alpha} \leq 1$, we obtain

$$m[n]^2c[n]\eta_2(n) \leq m[n]^2m[n-1]\mathfrak{C}^*\mathfrak{c}_2\Xi M_n^{4+\beta-l} \leq m[n]^2m[n-1]\Theta^*\Theta^{-1}M_n^{-\alpha} \leq \Theta^*,$$

which is exactly hypothesis (ii) of lemma 8.

The rest of the proof is devoted to verifying the assumptions made in lemma 8. To check (i) we need proposition 5. Hypothesis (i), commutativity of

ν and τ, was done in lemma 11. Hypothesis (ii) is exactly the statements (iv), (v) of lemma 10. Hypothesis (iii) coincides with statement (iii) of lemma 10 (which was achieved by the non linear transformation h). Finally, proposition 5 (iv) is exactly the statement (ii) of lemma 10. This altogether verifies hypothesis (i) of lemma 8. This and statements (iv) and (v) of lemma 10 (smallness of the l-th order erivatives of ν,τ) imply hypothesis (iii) and (iv). This completes the proof of proposition 7. □

The next lemma links the growth of the M_n to the Liouville exponent of ω.

Lemma 12: Let Δ_n and M_n be as in proposition 7. Then, if ω has Liouville exponent γ, i.e. $m[n] \le \mathfrak{K} M_n^{\gamma}$, for some constant $\mathfrak{K} < \infty$ (which depends on c[0]).

Proof. Define $\sigma_n = q[n]M_n^{-1}$, where $q[n+1] = m[n]q[n] + q[n-1]$, are the denominators in the convergents of ω. We derive inductively (recall $\Delta_n = (m[n] + \Delta_{n-1})$

$$\sigma_{n+1} = m[n]\sigma_n\Delta_n + \sigma_{n-1}\Delta_n\Delta_{n-1}$$

$$\le \left[(1 + \Delta_{n-1}/m[n])^{-1} + (1 + m[n]/\Delta_{n-1})^{-1}\right] \max(\sigma_n, \sigma_{n-1}),$$

that is $\sigma_{n+1} \le \max(\sigma_n, \sigma_{n-1})$. If we put $\sigma^* = \max(\sigma_1, \sigma_2)$, then $q[n] \le \sigma^* M_n$ for all $n \in \mathbb{N}$. It is known (cf. Hawkins - Schmidt [2]), that:

(i) $\mathfrak{b}_n m[n] \le \mathfrak{b}_{n-1} \le \mathfrak{b}_n(2 + m[n])$, where $\mathfrak{b}_n = |q[n] - \omega p[n]|$,

(ii) $\mathfrak{b}_{n-1} \le q[n]^{-1} \le \mathfrak{b}_{n-1}(1 + m[n]^{-1})$, for $n \ge 2$.

By definition of the Liouville exponent, for a positive C: $\mathfrak{b}_n \ge Cq[n]^{-1-\gamma}$. It is

$$\mathfrak{b}_n \ge Cq[n]^{-1-\gamma} \ge CM_n^{-1-\gamma}\sigma^{*-1-\gamma},$$

and on the other hand

$$\mathfrak{b}_n \le \mathfrak{b}_{n-1}m[n]^{-1} \le q[n]^{-1}m[n]^{-1} \le M_n^{-1}m[n]^{-1}.$$

The last step follows from the fact $M_n \le q[n]$. Hence $m[n] \le \mathfrak{K} M_n^{\gamma}$ with $\mathfrak{K} = C^{-1}\sigma^{*1+\gamma}$. □

Remark: We cite the convergence result of proposition 7 in the form

$$\eta(n)^{\gamma/(l-4-\beta)} \leq (\mathfrak{C}^*\eta(0))^{\gamma/(l-4-\beta)} M_n^{-\gamma},$$

and combine it with lemma 12, $m[n] \leq \mathfrak{M} M_n^{\gamma}$. Hence

$$m[n]\eta(n)^{\gamma/(l-4-\beta)} \leq \mathfrak{M}^*,$$

for all $n \in \mathbb{N}$, where $\mathfrak{M}^* = \mathfrak{M}(\mathfrak{C}^*\eta(0))^{-\gamma/(l-4-\beta)}$ is a constant. We shall use this formula in the next section frequently.

5. Invariant curves.

Lemma 9. The pair (U,T) possesses an invariant, Lipshitz continuous and transitive curve μ of rotation number ω.

Proof: We shall do the proof by 'pulling back' cycles, a method which was introcuded by D. Rand in [7]. First a definition: A **cycle** for (U,T) is an orbit segment $\mu = \{\mu[i] \in \mathbb{R}^2\colon\ i \in J\}$ under $\mathfrak{K}$ (see section 2 for the definiton of $\mathfrak{K}$) with $\mu[i+1] = \mathfrak{K}(\mu[i])$ for $1 \leq i < |\mu|$, where $\mathfrak{J} = \{1,2,\ldots,|\mu|\}$ is a numbering and $|\mu|$ is the lenth of the cycle (i.e. the number of points). As this definition suggests, a cycle resembles an invariant set when we look at some limited number of iterations of $\mathfrak{K}$. In the following we shall construct cycles with increasing length and the idea is, that they will converge to an invariant curve μ. We begin with a cycle for (U_n,T_n) which consists of two points. Out of this we construct a larger cycle for (U_{n-1}, T_{n-1}) and inductively for (U_i,T_i), for $i = n,\ldots,0$, the length of the cycles increases at least exponentially fast with i decreasing. We refer to this procedure as 'pulling back' of cycles. We always get eventually a cycle for (U,T). As we increase n, the length of the pulled back cycle for (U,T) will increase at least exponentially fast, and converges as a set to a lipshitz continuous curve μ, which is invariant under $\mathfrak{K}$.

Let L, R be the left respectively right halfplane in x,y - coordinates, and define $\mathfrak{a} = |\mu \cap L|$ and $\mathfrak{b} = |\mu \cap R|$. Then we shall say μ is of type $(\mathfrak{a},\mathfrak{b})$ for (U,T). In the following we count the y-axis $\{(x,y) \in \mathbb{R}^2\colon\ x = 0\}$ alternating to L and to R. If $N_m(U,T)$ has a cycle of type $(\mathfrak{a},\mathfrak{b})$, then (U,T) has one of type $(\mathfrak{b} + m\mathfrak{a}, \mathfrak{a})$, namely

$$\mu^{\sim} = \Lambda^{-1}\mu \cup \bigcup_{1 \leq u < m} U[u](\Lambda^{-1}\mu \cap R'),$$

as a set, where $R' = \Lambda^{-1}L$ plays the role of the right half plane. The notion of 'right' and 'left' half plane cannot be taken so literally as it was introduced in section 2, because the coordinate change Λ which has a non linear part maps the y-axis $Y = \{(x,y) \in \mathbb{R}^2\colon\ x = 0\}$ in general to some curve close to Y but not onto itself. In the sequel we will neglect this subtlety. There is a unique numbering that makes $\mu^{\sim}$ to a cycle. Introduce renormalisation indices and let μ_{n+1} be a cycle of (U_{n+1}, T_{n+1}) with Lipshitz constant l_{n+1}, then μ_n has Lipshitz constant l_n. Let $\Omega = \bigcap_{t\in\mathbb{N}} \Lambda_t^{-1}\ldots\Lambda_0^{-1}(\bullet)$, which is a single point because $\bullet$ is compact, and construct a sequence of pulled back cycles $\{\mu(i) \in \mathbb{R}^2 : i \in \mathbb{N}\}$ of (U_0, T_0). Put $\Omega_i = \Lambda_i\ldots\Lambda_0\Omega$. We start the construction of $\mu(i) = \mu_0(i)$ with $\mu_i(i) = \{T_i^{-1}\Omega_i, \Omega_i\}$, which is a two cycle. The Lipshitz constants of $\mu_i(i)$ are bounded by $\lambda m[i]\eta(i)$, for some constant λ (since the non linearities of U_i, T_i are estimated by such an expression), or, due to the remark to lemma 12, we have $l_i(i) \le \lambda\eta(i)^{1-\gamma/(1-4-\beta)}$ (replacing λ by $\lambda\Xi^*$). The constant λ shall be determined shortly. The cycle $\mu_i(i)$ $(i-k)$ - times pulled back will be denoted by $\mu_k(i)$, $k = i,\ldots,0$. Drop the index i, so that $l_k(i)$ now reads l_k and let us assume $l_k \le \lambda\eta(k)^{1-\gamma/(1-4-\beta)}$ holds true for $k = i,\ldots,n+1$, for some $n \le i$. For the induction step from $n+1$ to n the Lipshitz constants change as indicated in the following diagram.

$$l_{n+1} \xrightarrow{Dh^{-1}} \tilde{\Gamma}_n \xrightarrow{DA'^{-1}_n} \Gamma_n \xrightarrow{DU_n[u],\ 0\le u<m[n]} l_n.$$

For the first estimate we need that $h: (x,y) \to (x+F(x,y), y+G(x,y))$ (without introducing an index n) is near to the identity transformation, in particular $\|\partial F\|, \|\partial G\| \le (1+t^*)\Gamma_{y,n}\Gamma_{x,n}^{-1}\eta_2(n)$, because of the remark to lemma 10. Hence $\tilde{\Gamma}_n \le l_{n+1} + (1+t^*)\Gamma_{x,n}^{-1}\Gamma_{y,n}\eta_2(n)$, and for the second step one finds $\Gamma_n \le \Gamma_{x,n}\Gamma_{y,n}^{-1}\tilde{\Gamma}_n$. The two inequalities combined yield

$$\Gamma_n \le \Gamma_{x,n}\Gamma_{y,n}^{-1}l_{n+1} + (1+t^*)\eta_2(n) \le (g^{-1/4}\lambda + 1)\eta_2(n)^{1-\gamma/(1-4-\beta)},$$

(g is the golden mean) where we used that $\Gamma_{x,n}^{-1}\Gamma_{y,n}$ is less than $g^{-1/2}$ (for almost all $n \in \mathbb{N}$, except in that special case discussed in lemma 10) and the fact that t^* can be made arbitrary small by chosing Θ^* (i.e. Θ) small enough, here we made $t^* \le g^{-1/4} - g^{-1/2}$. For the last step in the diagram we get

$$\text{(5-1)} \qquad l_n \le \Gamma_n\big[1 - \big(m[n]c[n](1 + O^*(6\varepsilon_{2,m[n]}(n))) + 1\big)\Gamma_n\big]^{-1} + \varepsilon_{2,m[n]}(n).$$

(Under an affine shear $S: (x,y) \to (x+sy, y)$, $s > 0$, derivatives transform like $dy/dx \to (dy/dx)/(1 + s\,dy/dx)$.) For Θ and Ξ small enough (depending on λ, but λ does not depend on Θ and Ξ) we can achieve that

$$\big(m[n]c[n](1 + 6\varepsilon_{2,m[n]}(n)) + 1\big)\Gamma_n \le 1/6,$$

that is, the denominator in (5-1) is bigger than 5/6. Expand (5-1) into a geometric series expansion cut off after the second term. Hence

$$l_n \le \Gamma_n\big[1 + (6/5)\Gamma_n\big(m[n]c[n](1 + 6\varepsilon_{2,m[n]}(n)) + 1\big)\big] + \varepsilon_{2,m[n]}(n),$$

For Θ and Ξ small enough, (assuming $\lambda \geq \mathfrak{z}^{3/4}$) the coefficient of Γ_n^2 is in modulus less than

$$2\lambda(6/5)(m[n]c[n](1+6\Theta^*)+1)\eta(n)^{1-\gamma/(l-4-\beta)} \leq \sqrt[4]{\mathfrak{z}}-1,$$

using $\varepsilon_{2,m[n]}(n) \leq 2m[n]\eta_2(n) \leq \Theta^*$ (because $1+\mathcal{O}_p((m+1)^2c\eta) \leq 2$) for Θ small enough). The same argument in conjunction the remark made to lemma 12 provides $\varepsilon_{2,m[n]}(n) \leq 2\Xi^*\varepsilon_2^{\gamma/(l-4)}$, and furthermore

$$l_n \leq [\sqrt[4]{\mathfrak{z}}(\mathfrak{z}^{-3/4}+\lambda^{-1})\lambda + 2\Xi^*\varepsilon_2^{\gamma/(l-4-\beta)}]\eta_2(n)^{1-\gamma/(l-4-\beta)}.$$

For λ large enough the factor $\sqrt[4]{\mathfrak{z}}(\mathfrak{z}^{-1/4}+\lambda^{-1})$ is less than $\mathfrak{z}^{-1/8}$ and furthermore $\mathfrak{z}^{-1/8}+2\lambda^{-1}\Xi^*\varepsilon_2^{\gamma/(l-4)} \leq 1$. This determines the constant λ and proves that the Lipshitz constants l_i are bounded uniformly for $i \to \infty$. The pulled back cycles $\mu(i)$ as sets form a nested sequence: $\{\mu(0)\} \subset \ldots \subset \{\mu(i)\} \subset \{\mu(i+1)\} \subset \ldots$. It remains to show that the limit set $\mu = \bigcup_{i\leq 0}\{\mu(i)\}$ is transitive, i.e. that the closure of $\pi_x\mu$ is an interval in $\mathbb{R}$. Define

$$k_n(i) = \max\{|\pi_x\xi-\pi_x\zeta| : (\xi,\zeta) \in \mu_n(i)\times\mu_n(i),\ \pi_x\xi \text{ and } \pi_x\zeta \text{ are neighbours on } \mathbb{R}\}.$$

Since all the Lipshitz constants are uniformly bounded, to prove that there are no 'gaps' in μ, it is sufficient to show that $k_n(i)$ goes to zero for $i \to \infty$. Drop the index i, and we deduce for $k_{n+1} \to k_n$, $0 \leq n < i$,

$$k_n \leq k_{n+1}\Gamma_{x,n}^{-1}[1+2\Gamma_n(m[n]c[n]+1)]$$

$$k_n \leq k_{n+1}\Gamma_{x,n}^{-1}[1+2(\mathfrak{z}^{-1/4}\lambda+1)(m[n]c[n]+1)\eta_2(n)^{1-\gamma/(l-4-\beta)}]$$

The second factor in the squared brackets is less than $\mathfrak{z}^{1/4}-1$, say, provided Θ is small enough (so that $2(\mathfrak{z}^{-1/4}\lambda+1)\Theta^* \leq \mathfrak{z}^{1/4}-1$, substituting $mc\eta_2(n)^{1-\gamma/(l-4-\beta)}$ by Θ^*, where we used the remark to lemma 12). Hence $k_n \leq k_{n+1}\mathfrak{z}^{-1/4}$, since $\Gamma_{x,n}^{-1}$ is (almost always) less than $\mathfrak{z}^{-1/2}$. Hence $k_0 \leq k_i\mathfrak{z}^{-i/4}$ converges at least exponentially fast to zero for $i \to \infty$, since the starting k_i are bounded by 2, say. This concludes the proof that μ is transitive. That μ has under $\mathbf{F}$ rotation number ω is clear from the construction. □

Fix some renormalisation index n, set $(U,T) = (U_n,T_n)$, $c = c[n]$, $m = m[n]$, $\eta = \eta(n)$ and h for the non-linear part of the coordinate transformation Λ_n. Before moving on to the construction of the smooth curves we shall prove the next two lemmas.

Lemma 14: Let $y(x)$ be a ψ-times differentiable, real valued function on some interval $J \subset \mathbb{R}$. Set $\gamma = \max_{1\le r\le \psi} (\|\partial_x{}^r y(x)\|_J$ and let $y' = y \circ h^{-1}$, where $h = h_{l'} \circ \ldots \circ h_0$ (l' is either l or $l+1$) is the non linear coordinate transform introduced in lemma 10. Let $\upsilon[i]$ be the generating function of h_i, $i = 0,\ldots,l'$. Then there are constants $\theta^*, \gamma_0 > 0$ so that

$$\|\partial_x{}^r y'\|_J \le \gamma^{1/4}(\gamma + \|\partial^{r+1}\upsilon[0]\| + \ldots + \|\partial^{r+1}\upsilon[l']\|),$$

for $r = 1,\ldots,\psi$, whenever $\|\partial^{r+1}\upsilon[0]\| + \ldots + \|\partial^{r+1}\upsilon[l']\| \le \theta^*$, $r = 1,\ldots,\psi$, and $\gamma \le \gamma_0$.

Proof. We shall prove this lemma separatly for each of the transformations h_i, $i = 0,\ldots,l'$, which together made h. Let $\upsilon(x,x')$ be the generating function of h_0.

We use the chain and product rules for differentiation and get a for $\partial_x{}^r y'(x)$ an expression similar to the one encountered in (4-8) of lemma 8. Recall that $h: (x,y) \to (x',y')$, where $x' = \pi_x U(x,y)$ and $(y,y') = (-\upsilon_{10}(x,x'), \upsilon_{01}(x,x'))$. Hence (where $\mathbf{B}(q,t)$ stands for the binomial coefficient $q!/((q-t)!t!)$)

$$(5\text{-}2) \quad \partial_x{}^r y' = \partial^r y + \sum_{0\le q\le r} \sum_{0\le s\le r-q} \sum_{0\le t\le q} \mathbf{B}(q,t)\upsilon_{q-t+1,t+s}((\partial_y x')\partial y)^t \mathbf{P}_r{}^{q,s}$$

where $\mathbf{P}_r{}^{q,s}$ are sums of monomials of the form $\prod_{1\le b\le s} d_x{}^{i[b]}(\partial_y x')\partial y$, for integers $i[b] \ge 2$ for which $\sum_{1\le b\le s} i[b] = r-q$. The implicit function theorem applied to $y = -\nu_{10}(x,x'(x,y))$ yields $d_y x' = -\nu_{11}(x,x')^{-1}$. Furthermore, by Lagranges' formula,

$$d_x{}^{i[b]}(\partial_y x')\partial y = \sum_{0\le w\le i[b]} \mathbf{B}(i[b],w)(d_x{}^w(\partial_y x'))(\partial^{i[b]-w+1} y).$$

If $w \ge 1$ it follows by the chain rule that there is at least one derivative $\partial^{w^*} y$, $w^* \ge 1$, which is multiplied with $\partial^{i[b]-w+1} y$. This means that $t = 0$ is the only summand that is 'linear' in ∂y. Hence there exist polynomials $\mathbf{P}_{s,b}(y)$, $\mathbf{P}_{s,b}(0) = 0$ with positive coefficients, so that

$$\|d_x{}^{i[b]}(\partial_y x')\partial y\| \le \gamma(1 + \mathbf{P}_{s,b}(\gamma)),$$

for $b = 1,\ldots,s$. Observe in (5-2) that $t = 0$ is the only term which is 'linear in $\partial y'$. Furthermore, $\mathbf{P}_r{}^{q,s}$ is trivial ($= 1$) only if $s = 0$, but this is possible only if $q = r$. Hence there is a polynomial $\mathbf{P}^*(\gamma)$ with positive coefficients and which begins with quadratic terms in γ, so that

$$\|\partial_x{}^r y'\| \le \|\partial^r y\| + \|\upsilon_{r+1,0}\| + \sum_{0\le q\le r} \sum_{1\le s\le r-q} (\|\partial^{q+s+1}\upsilon\|)\gamma^s \prod_{1\le b\le s} (1 + \mathbf{P}_{s,b}(\gamma)) + \mathbf{P}^*(\gamma).$$

Given some $x > 1$, then there exist positive constants θ^*, γ_0 so that

$$\|\partial_x{}^r y'\| \le x\gamma + \|\upsilon_{r+1,0}\|$$

whenever $\|\partial^{q+s+1}\upsilon\| \le \theta^*/l'$ for $q+s = 0,\dots,r$ and $\gamma \le \gamma_0$. Now choose positive θ^*, γ_0 so that $x = \gamma^{1/4l'}$. We repeat the same argument for h_1 instead of h_0, and so on. This concludes the proof of the lemma. □

Lemma 15: Let $y^\sim(x)$ be a real valued, j-times differential function defined on some interval $J^\sim \subset \mathbb{R}$, and let $\nu[u]$ be the generating function of some $U[u]$. Set $\gamma^\sim = \max_{r=1,\dots,j} \|\partial_x{}^r y^\sim\|_{J^\sim}$ and suppose $mc\gamma^\sim \le \gamma^\sim{}_0$ for some small constant $\gamma^\sim{}_0 > 0$. Then

$$\|\partial_x g[u]\| \le 1 + \theta' u^2 c\eta + uc\gamma^\sim + \mathcal{P}(\theta' u^2 c\eta + uc\gamma^\sim),$$
$$\|\partial_x{}^r g[u]\| \le \qquad \theta' u^2 c\eta + uc\gamma^\sim + \mathcal{P}(\theta' u^2 c\eta + uc\gamma^\sim),$$

for $r = 2,\dots,j$, $u = 1,\dots,m$ and a positive constant θ', where $\mathcal{P}(\gamma)$ is a polynomial in γ without constant and linear terms, and where $x = g[u](x')$ is the function determined by the implicit function theorem out of $y^\sim(x) = -\nu_{10}[u](x,x')$.

Proof. Recall that up to some constant $\nu[u] = (x-x'-1)^2/(2+2uc)$ + remainder. The proof is by induction on r. We shall begin with the first derivative. In that case we get (write ∂ for the differentiation if there is only one argument)

$$\begin{aligned}\partial g[u](x) &= -\nu_{11}[u](x,x')[\nu_{20}[u](x,x') + \partial y^\sim(x)]^{-1}\\ &= [(uc+1)^{-1} + O^*(9(u+1)\eta_2)]\cdot[(uc+1)^{-1} + O^*(9(u+1)\eta_2) + \partial y^\sim(x)]^{-1}\\ &= 1 + \mathcal{P}(\theta' u^2 c\eta + uc\gamma^\sim).\end{aligned}$$

for a constant θ', where $\mathcal{P}(y)$ is a positive polynomial of the form $\gamma + c_2\gamma^2 + c_3\gamma^3 + \dots$, without constant term and where the linear term has coefficient 1. In the last estimate we used $\varepsilon_{2,u} \le 9(u+1)\eta_2$, which we obtained in lemma 8. By way of differentiating $\partial g[u]$ once more, $r = 2$, we obtain explicitly:

$$\begin{aligned}\partial^2 g[u] = &-[\nu_{21}[u]\partial g[u] + \nu_{12}[u]]\cdot[\nu_{20}[u] + \partial y^\sim]^{-1}\\ &- \nu_{11}[u][(\nu_{30}[u] + \partial^2 y^\sim)\partial g[u] + \nu_{21}[u]]\cdot[\nu_{20}[u] + \partial y^\sim]^{-2}\\ = &\ \mathcal{P}(\gamma^\sim(uc+1)) + (uc)^{-1}O(u\eta)[(uc)^{-1} + \mathcal{P}(\gamma^\sim(uc+1))]^2 = \mathcal{P}(\gamma^\sim(uc+1)),\end{aligned}$$

for $u = 1,\dots,m$, where we used that $\partial g[u] = 1 + \mathfrak{P}(\theta' u^2 c\eta + uc\boldsymbol{\gamma}^{\sim})$. As doen here we shall in the sequel drop the arguments (note that ∂_x on $\nu[u]$ acts like $(\partial g[u])\partial_1 + \partial_2$). For simlicity we use the same symbol $\mathfrak{P}(\boldsymbol{\gamma})$ for polynomials that are in fact different but are all of the same form, the one given in the last paragraph. For the step from $r-1$ to r we obtain inductively:

$$(5\text{-}3) \qquad \partial^r g[u] = -\sum_{0\le s\le r-1} \mathfrak{P}(r-1,s)[\partial^{r-1-s}\nu_{11}[u]]\partial^s[\nu_{20}[u] + \partial y^{\sim}]^{-1}$$

In case $r-1-s > 0$, the first factor involves at least third order derivatives and is henceforth $O(u\eta)$ (by virtue of lemma 8). If $r-1-s = 0$ we use the approximation $\nu_{11}[u] = -(1+uc)^{-1} + O(u\eta)$. The second factors in (5-3) decompose into

$$\partial^s[\nu_{20}[u] + \partial y^{\sim}]^{-1} = \sum_{1\le q\le s} [\nu_{20}[u] + \partial y^{\sim}]^{-q-1} \prod_{1\le t\le q} \partial^{s[t]}[\nu_{20}[u] + \partial y^{\sim}],$$

where the integers $s[t]$ are at least one and $\sum_{1\le t\le q} s[t] = s$. Using the fact $\nu_{20}[u] = (1+uc)^{-1} + O(u\eta)$, for $u = 1,\dots,m$, the denominator is estimated by

$$[\nu_{20}[u] + \partial y^{\sim}]^{-q-1} = [uc(1 + O(u^2c\eta) + \mathfrak{P}(\boldsymbol{\gamma}^{\sim}))]^{q+1} = O(uc)^{q+1},$$

and for the product we derive

$$\Big\|\prod_{1\le t\le q} \partial^{s[t]}[\nu_{20}[u] + \partial y^{\sim}]\Big\| \le \prod_{1\le t\le q} [\theta' u\eta + \boldsymbol{\gamma}^{\sim}]$$

$$= (\theta' u\eta + \boldsymbol{\gamma}^{\sim})^q,$$

for a constant θ'. We collect the error-terms from the last two inequalities and find that the resulting term is of size $O(uc(\theta' u^2 c\eta + uc\boldsymbol{\gamma}^{\sim})^q)$ for $q = 1,\dots$. The second factor on the right hand side of (5-3) therefore is of size $O(\theta' u^3 c^2\eta + u^2c^2\boldsymbol{\gamma}^{\sim})$, whereas (5-3) itself is to estimate like

$$\|\partial^r g[u]\| = [(1+uc)^{-1} + O(u\eta)]O(\theta' u^3c^2\eta + u^2c^2\boldsymbol{\gamma}^{\sim}) = \mathfrak{P}(\theta' u^2 c\eta + uc\boldsymbol{\gamma}^{\sim}),$$

for $u = 1,\dots,m$, where $\mathfrak{P}$ is a polynomial of the same kind as those above. Here we used the assumption that $uc\boldsymbol{\gamma}^{\sim}$ is bounded by $\boldsymbol{\gamma}^{\sim}_0$. This proves the lemma.

□

After these preparations, we shall prove the main result of this section.

Proposition 16: The invariant curve μ is a graph of class $C^{j+\text{Lipshitz}}$, for $j \le \psi - 1 = [l-4-\beta-\gamma]$.

Proof. We extend the notion of 'pulling back' to curves that interpolate cycles similar to those constructed in lemma 13. (The interpolated cycles have to chosen a bit different from those in lemma 13. This is to get uniform bounds on the higher derivatives of the initiating curves. We shall show at the end of the proof, that the pulled back curves converge to μ.) The curves will not have an invariance property apart from the fact that each contains a set which is a cycle. The idea is, to show that the pulled back curves $\chi(i)$ are uniformly smooth for all derivatives up to $\psi = [l-3-\beta-\gamma]$. By the theorem of Arzela - Ascoli the sequence (or a subsequence, to be more precise) converges uniformly in $C^{\psi-1}$ to a limit χ, which here coincides with μ.

Fix some index $i \in \mathbb{N}$ and set $\Omega^*_i = (0, -\partial_1(\nu \oplus \tau)(0,0))$, where (ν,τ) are the generating functions of (U_i,T_i). Unlike to the situation in lemma 13 we begin the pulling back with the two-cycle $\{T_i^{-1}\Omega^*_i, \Omega^*_i\}$. Furthermore, to this set we add the point $U_i^{-1}\Omega^*_i$. These three points do not form a cycle but they will grow up to one, since we identify

$$U_{i-1}[m[i-1]]\Lambda_{i-1}^{-1}T_i^{-1}\Omega^*_i = \Lambda_{i-1}^{-1}U_i^{-1}\Omega^*_i.$$

Note, that lengths of the intervals $(\pi_x T_i^{-1}\Omega^*_i, \pi_x\Omega^*_i)$, $(\pi_x\Omega^*_i, \pi_x U_i^{-1}\Omega^*_i)$ uniformly in n are bounded from above and below. This gives readily uniform bounds on the derivatives of the curves interpolating the initial points $\{T_i^{-1}\Omega^*_i, \Omega^*_i, U_i^{-1}\Omega^*_i\}$. We connect these three points by a C^∞-curve χ_i (which stands for $\chi_i(i)$), so that

$(i)_i$ $U_i\chi_i$ and $T_i\chi_i$ join C^l-up in the point Ω^*_i;

$(ii)_i$ $U_i\chi_i$ and χ_i join up in Ω^*_i;

$(iii)_i$ $U_i\chi_i$ and χ_i join up in Ω^*_i.

It is clear, that the conditions $(ii)_i$ and $(iii)_i$ imply $(i)_i$.

Let us demonstrate the induction step $n-1 \to n$. Suppose we got the curves χ_k for $k = n-1,\dots,i$. Put $\chi_{n+1} = \{y_{n+1}(x)\colon x \in J\}$ for some interval J. The coordinate transformation Λ_n brings χ_{n+1} to $\chi^\sim_n = \Lambda_n \circ \chi_{n+1} \circ \Lambda_n^{-1} = \{(x, y^\sim_n(x)) \in \mathbb{R}^2\colon x \in J^\sim\}$, for the new interval $J^\sim = \Gamma_{x,n}^{-1}J$. The new curve χ_n is pieced together from copies of $\chi^\sim_n$ as follows:

$$\chi_n = \chi^\sim_n \cup \bigcup_{0 \le u < m[n]} U_n[u](\chi^\sim_n \cap R'),$$

where R' plays the role of the right half plane, i.e. $R' = \Lambda^{-1}L$, where L represents the 'left half of the plane' one renormalisation step before. Having applied the transformation Λ_n^{-1} the joining up condition $(iii)_{n+1}$ reads: $T_n\chi^\sim_n$ and $\chi^\sim_n$ join up at $\Omega^*_n = \Lambda_n^{-1}(\Omega^*_{n+1})$. The points $U_n[u+1](\Omega^*_n)$, $U_n[u](T_n\Omega^*_n)$ are to be identified because $U_n[u+1] = T_n \circ U_n[u]$. Furthermore, by $(ii)_{n+1}$ it follows that $U_n[u+1]\chi^\sim_n$ and $U_n[u]\chi^\sim_n$ join together at the point $U_n[u](\Omega^*_n)$. Hence χ_n is a smooth curve in C^l.

To complete the induction, one has to verify the joining up conditions hold true for χ_n.

$(iii)_n$ By condition $(iii)_{n+1}$ it follows that $T_n\chi^\sim_n$ and $\chi^\sim_n$ join up at Ω^*_n, since $U_{n+1} = \Lambda_n \circ T_n \circ \Lambda_n^{-1}$;

$(ii)_n$ By the construction of the renormalised pair it is clear that all iterates T_n^k are well defined for k running up to $m[n]$. Thus, repeatedly applying $(iii)_n$ we get that $T_n^{m[n]}\chi^\sim_n$ and $\chi^\sim_n$ join up at Ω^*_n. From condition $(ii)_{n+1}$ it is known that $T_n^{m[n]}U_n\chi^\sim_n$ and $\chi^\sim_n$ join up at Ω^*_n, and therefore $U_n\chi^\sim_n$ and $\chi^\sim_n$ join up at Ω^*_n.

$(i)_n$ As we pointed out above, this condition follows from $(iii)_n$ and $(ii)_n$.

In the following we shall estimate the derivatives of the χ_n. Set $\mathfrak{u}_n[j] = \|\partial_x^j y_n(x)\|_J$, for $j = 1,\dots,\psi$. As indicated in the diagram the $\mathfrak{u}$'s are affected by the various operations we apply successively to χ_{n+1}:

$$\mathfrak{u}_{n+1}[j] \xrightarrow{Dh^{-1}} \mathfrak{u}^\wedge_n[j] \xrightarrow{D\Lambda'^{-1}_n} \mathfrak{u}^\sim_n[j] \xrightarrow{DU_n[u],\ 0\le u<m[n]} \mathfrak{u}_n[j].$$

For the first step we prepared lemma 14. Let us assume $\mathfrak{u}_{n+1}[j] \le \mathfrak{z}_0$, and the coordinate transform is close enough to the idendity, i.e. $\|\upsilon_{q,s}[0]\| + \dots + \|\upsilon_{q,s}[l']\| \le \theta^*$ for $q+s = 1,\dots,j+1$ (where l' is the number of transformations (either l or $l+1$) into which we split h in lemma 9). Hence

$$\mathfrak{u}^\wedge_n[j] \le \mathfrak{z}^{1/4}(\mathfrak{u}_{n+1}[j] + \|\upsilon_{j+1,0}[0]\| + \dots + \|\upsilon_{j+1,0}[l']\|) \tag{5-4}$$

The second step in the diagram involves the affine coordinate transform Λ'_n:

$$\mathfrak{u}^\sim_n[j] \le \mathfrak{u}^\wedge_n[j]\Gamma_{x,n}{}^j\Gamma_{y,n}{}^{-1}. \tag{5-5}$$

We shall proof that $\mathfrak{u}_n[j] \le \mathfrak{C}\eta(n)^{1-\gamma/(1-4-\beta)}$, for some constant $\mathfrak{C} > 0$ whichshall be determined below. To begin with, we note that the distances between the initiating points $T_i^{-1}\Omega^*_i$, Ω^*_i, $U_i^{-1}\Omega^*_i$, are by construction bounded from below (and above) uniformly in i. It is therefore clear that $\mathfrak{u}_i[j] \le \mathfrak{C}\eta(n)^{1-\gamma/(1-4-\beta)}$, for a positive $\mathfrak{C}$ and all $j = 1,\dots,\psi$, $i \in \mathbb{N}$. This combined with (5-4) and (5-5) yields

$$\begin{aligned}\mathfrak{u}^\sim_n[j] &\le \Gamma_{x,n}{}^j\Gamma_{y,n}{}^{-1}\mathfrak{z}^{1/4}(\mathfrak{C}\eta(n+1)^{1-\gamma/(1-4-\beta)} + \|\upsilon_{j+1,0}[0]\| + \dots + \|\upsilon_{j+1,0}[l']\|)\\ &\le \Gamma_{x,n}{}^j\Gamma_{y,n}{}^{-1}\mathfrak{C}\mathfrak{z}^{1/4}\eta(n+1)^{1-\gamma/(1-4-\beta)} + (l+2)\eta_{j+1}(n),\end{aligned}$$

where we used $\|\upsilon_{j+1,0}[i]\| \le \Gamma_{x,n}{}^{-j}\Gamma_{y,n}\eta_{j+1}(n)$, $i = 0,\dots,l'$, as was pointed out in the remark to lemma 10. Using the convergence result of proposition 7 in the form $\eta(n+1) \le \eta(n)\Delta_n{}^{1-4-\beta}$ one finds (observe that $\gamma/(1-4-\beta) < 1/2$)

$$\mathfrak{t}^{\sim}_n[j] \le C D_n \delta^{1/4} \Delta_{n-1} \Delta_n^{l-j-3-\beta-\gamma} \eta(n)^{1-\gamma/(l-4-\beta)} + (l+2)\eta_{j+1}(n)$$

(5-6) $$\le C D_n \delta^{1/4} \Delta_{n-1} \eta(n)^{1-\gamma/(l-4-\beta)} + (l+2)\eta_{j+1}(n),$$

for constants $D_n \le (1 + \kappa m[n]^2 m[n-1]\eta_2)$ to level out the error we made by replacing $\Gamma_{x,n}$ by Δ_n^{-1} and $\Gamma_{y,n}$ by $\Delta_n^{-1}\Delta_{n-1}^{-1}$, where $\kappa > 0$ is some fixed constant. We used furthermore the assumption $j \le l-3-\beta-\gamma$, i.e. $\Delta_n^{l-j-3-\beta-\gamma} \le 1$. We multiply (5-6) on both sides with mc. Hence

$$mc\mathfrak{t}^{\sim}_n[j] \le C D_n \delta^{1/4} \Delta_{n-1}\, cm\eta(n)^{1-\gamma/(l-4-\beta)} + (l+2)mc\eta_{j+1}(n)$$

$$\le C D_n \delta^{1/4} \eta(n)^{1-2\gamma/(l-4-\beta)} + \varepsilon^{\sim}_0/2,$$

where we used that $(l+2)cm\eta_{j+1}(n) \le \varepsilon^{\sim}_0/2$ if Θ^* of lemma 8 is small enough ($\Theta^* \le \varepsilon^{\sim}_0/(2(l+2)\kappa_{j+1})$), and where we replaced $\Delta_{n-1}c$ by 1. The small error, $|\log \Delta_{n-1}c| \le \kappa m^2 c\eta_2$, for some constant $\kappa > 0$, is swallowed by the D_n, where κ will increase slightly. Let D be an upper bound for the D_n. Since $2\gamma < l-4-\beta$, by shrinking Ξ, we can achieve that $\eta(n)^{1-2\gamma/(l-4-\beta)} \le (C D \delta^{1/4})^{-1}\varepsilon^{\sim}_0/2$, whenever $\eta(0) \le \Xi$. Thus

$$mc\mathfrak{t}^{\sim}_n[j] \le \varepsilon^{\sim}_0.$$

It is now the turn of $\mathfrak{t}_n[j]$ to get estimated. In the following we will drop the renormalisation index n. Set $\chi = \chi_n$ and let (ν,τ) the generating functions of $(U,T) = (U_n,T_n)$. We call $\chi^*[u] \subset \chi$ the piece of curve that is created using $U_n[u]$:

$$\chi^*[u] = U_n[u](\chi^{\sim} \cap R) = \{(x, y^*[u](x)) \in \mathbb{R}^2 :\ x \in J^*[u]\}$$

for the interval $J^*[u] = \pi_x \chi^*[u]$. (Remember $\chi = (\chi^{\sim} \cap R) \cup \bigcup_{0 \le u < m[n]} \chi^*[u]$). We shall use generating functions to express $y^*[u](x)$, that is

$$y^*[u](x') = \nu_{01}[u](x,x'),$$

where $(x,x') \in J^{\sim} \times J^*[u]$. The first coordinate $x = x(x')$ has to be determined from

$$y^{\sim}(x) = -\nu_{10}[u](x,x').$$

By the implicit function theorem there exist functions $g[u]$: $x = g[u](x')$ for $x' \in J^*[u]$. The derivatives of $y^*[u](x) = \nu_{01}[u](x,x')$ lead to an expression similar to (4-8) of lemma 8 and to lemma 14, namely

$$\partial^j y^*[u] = \sum_{0 \le q \le j} \sum_{0 \le s \le j-q} \sum_{0 \le t \le q} B(q,t)\,\nu_{s+t,q-t+1}(\partial g[u])^t P_j^{q,s}[u],$$

where $\mathcal{P}_j^{q,s}[u]$ are sums of monomials of the form $\prod_{1\le b\le s} \partial^{i[b]}g[u]$, for integers $i[b] \ge 2$ and $\sum_{1\le b\le s} i[b] = j-q$. It was shown above that $m c \mathfrak{t}^{\sim}_n[r] = mc\|\partial^r y^{\sim}(x)\|_J{\sim} \le \mathfrak{Y}^{\sim}_0$ for $1 \le r \le \psi = [l-3-\beta-\gamma]$. We may therefore apply lemma 15. Observe

(i) $s = 0$ and $q < j$ is impossible, because all j derivatives must act on ν if $\mathcal{P}_j^{q,s}[u]$ happens to be trivial.

(ii) $(q,s) = (0,1)$ has the ν-derivative $\nu_{11}[u] = (uc)^{-1} + O(u\eta)$, which is multiplied by a factor $\mathcal{P}_j^{q,s}[u] = \partial^j g[u] = \theta' u^2 c\eta + uc\mathfrak{Y}^{\sim} + \mathfrak{B}(\theta' u^2 c\eta + mc\mathfrak{Y}^{\sim})$ and

(iii) all the others have $\mathcal{P}_j^{q,s}[u] = O(\theta' u^2 c\eta + uc\mathfrak{Y}^{\sim} + \mathfrak{B}(\theta' u^2 c\eta + uc\mathfrak{Y}^{\sim}))$.

Consequently

$$\|\partial^j y^*[u]\|_{J^*} = [(uc)^{-1} + O(u\eta)]O(\theta' u^2 c\eta + uc\mathfrak{Y}^{\sim} + \mathfrak{B}(\theta' u^2 c\eta + uc\mathfrak{Y}^{\sim})) \le \mathfrak{C}^{\wedge} m\eta$$

for all $u \le m$ and a constant $\mathfrak{C}^{\wedge} > 0$. Stressing the dependence on the iteration index n, this gives $\mathfrak{t}_n[j] \le \mathfrak{C}^{\wedge} m\eta$, $(m[n] = m)$, provided the hypothesis of lemma 15 is satisfied. As pointed out in the remark to lemma 12 we have $m[n]\eta(n)^{\gamma/(l-4-\beta)} \le \mathfrak{X}^*$ and therefore

$$\mathfrak{t}_n[j] \le \mathfrak{C}^{\wedge}\mathfrak{X}^*\eta(n)^{1-\gamma/(l-4-\beta)},$$

for all $u = 1,\dots,m[n]$. In particular, set $\mathfrak{C} = \mathfrak{C}^{\wedge}\mathfrak{X}^*$; and the induction hypothesis, $\mathfrak{t}_n[j] \le \mathfrak{C}\eta(n)^{1-\gamma/(l-4-\beta)}$ is satisfied.

Finally, convergence in C^0 has to be verified, since our initiating points Ω^*_i differ from Ω_i, the ones used in the lemma 13. We had to do that different choice to get uniform boundes on the derivatives of the smooth curves between $\pi_x T_i^{-1}\Omega^*_i$, $\pi_x\Omega^*_i$ and $\pi_x U_i^{-1}\Omega^*_i$. The distances are uniformly bounded from below (and above). This is not any more true if we replace Ω^*_i by Ω_i. Define

$$\mathfrak{d}_k(i) = \sup\{\inf\{|\xi-\zeta|: \zeta \in \Lambda_k\dots\Lambda_0\mu\}: \xi \in \chi_k(i)\},$$

for $k = 0,\dots,i$ and $i \in \mathbb{N}$, where $|\cdot|$ is the usual $\mathbb{R}^2$ metric. From the remark to lemma 10 it is known that $\|\partial^2 \upsilon[i]\| \le \Gamma_{x,k}^{-1}\Gamma_{y,k}\eta(k)$, $\forall\ i = 0,\dots,l'$. Assume that Θ and Ξ are small enough so that $(1 + 2\varepsilon_{2,m[k]}(k))(1 + \|\partial^2\upsilon[0]\| + \dots + \|\partial^2\upsilon[l']\|)$ is at most $\sqrt{\mathfrak{z}}$. Clearly, $\mathfrak{d}_i(i) \le 1$ for all $i \in \mathbb{N}$, and furthermore we deduce

$$\mathfrak{d}_k(i) \le (1 + 2\varepsilon_{2,m[k]}(k))\Gamma_{y,k}^{-1}(1 + \|\partial^2\upsilon[0]\| + \dots + \|\partial^2\upsilon[l']\|)\mathfrak{d}_{k+1}(i) \le \sqrt{\mathfrak{z}}\,\mathfrak{d}_{k+1}(i),$$

where we used that $\Gamma_{y,k}^{-1} \le \mathfrak{z}^{-1}$ (almost always). Hence, $\mathfrak{d}_0(i)$ decreases exponentially fast for $i \to \infty$. That proves that the pulled back curves $\chi(i)$ converge in C^{Lipshitz} to μ. □

Lemma 17: The map $\mathfrak{F}$ acts on μ $C^{j+\text{Lipshitz}}$-conjugate to $\mathfrak{R}_\omega$: $x \to x+\omega$ mod 1 on $\mathbb{R}/\mathbb{Z}$, for all $j \le \psi-1 = [L-3-\beta-\gamma] = [l-4-\beta-\gamma]$.

Proof. The proof goes parallel to proposition 16. Where appropriate we shall refer to the previous proposition and will not go into much details. For every $i \in \mathbb{N}$ we will construct a map $\mathfrak{h}(i)$: $\pi_x\chi(i) \to I = [0,1]$. As before, the $\mathfrak{h}(i)$, $i \in \mathbb{N}$, are bounded in C^ψ and converge to μ in C^0. Thus $\mathfrak{h}(i)$ converges for $i \to \infty$ in the $C^{\psi-1}$-topology uniformly to some $\mathfrak{h}$: $\pi_x\mu \to I = [0,1]$. If we set $\mathfrak{H} = \mathfrak{h}\circ\pi_x$, $\mathfrak{H}^{-1} = y\circ\mathfrak{h}^{-1}$, then $\mathfrak{H}$ is the transformation for which $\mathfrak{F}$ is conjugated to $\mathfrak{R}_\omega$, i.e. $\mathfrak{H}\circ\mathfrak{F}\circ\mathfrak{H}^{-1} = \mathfrak{R}_\omega$.

We start the construction of $\mathfrak{h}(i)$ by choosing a $\mathfrak{h}_i(i)$: $\pi_x\chi_i(i) \to I_i(i) = [0,1]$, which is ψ-times differentiable (or C^∞) and satisfies the three glueing up conditions of the previous lemma $(i)_i$, $(ii)_i$ and $(iii)_i$. Suppose we got $\mathfrak{h}_k$ for $n < k \le i$, then $\mathfrak{h}_n$ is constructed as follows:

$$\begin{array}{ccc} & DA_n^{-1} & \\ I_{n+1} & \longrightarrow\!\!\!\longrightarrow\!\!\!\longrightarrow & I^\sim_n = -\Gamma_{x,n}^{-1} I_{n+1}, \\ h_{n+1} & \longrightarrow\!\!\!\longrightarrow\!\!\!\longrightarrow & h^\sim_n. \end{array}$$

The maps on I_n corresponding to $U_{n[u]}$ on χ_n are translations

$$\begin{array}{llll} V_n[u]: & I^\sim_n & \to & I^*_n[u] = I^\sim_n - |I^\sim_n \cap L| - |I^\sim_n \cap R|(m[n]-u), \\ & x & \to & x - |I^\sim_n \cap L| - |I^\sim_n \cap R|(m[n]-u) \end{array}$$

for $0 \le u < m[n]$, where $|.|$ means the length of the interval. The intervals $I^*_n[u] = I^*_n[u](i)$ are uniformly in 'all indices' contained in a compact subset of $\mathbb{R}$. Define

$$\begin{array}{ll} \mathfrak{h}_n(x) = (V_n[u]\circ\mathfrak{h}^\sim_n\circ U_n[u]^{-1})(x, y_n(x)) & \text{for } x \in J^*_n[u] = \pi_x\chi^*_n[u] \\ \mathfrak{h}_n(x) = h^\sim_n(x) & \text{for } x \in I^\sim_n. \end{array}$$

This defines $\mathfrak{h}_n(x)$ on the interval $I_n = I^\sim_n \cup \bigcup_{0\le u<m[n]} I^*_n[u]$. Put $\mathfrak{h}_{\infty,c}$: $\mathbb{R}^2 \to \mathbb{R}^2$ for the linear map with slope c and constant part 0. For h_n determine the parameter c which minimises $\mathfrak{m}_n[0] = \|\partial(\mathfrak{h}_n - \mathfrak{h}_{\infty,c})\|$, where the norm is on I_n. In the same way we set $\mathfrak{m}_n[j]$ for $\|\partial^j(\mathfrak{h}_n - \mathfrak{h}_{\infty,c})\|$. Then the diagram holds

$$\begin{array}{ccccccc} & Dh^{-1} & & DA'^{-1} & & D[\mathfrak{h}^\sim_n \to V_n[u]\circ\mathfrak{h}^\sim\circ U_n[u]^{-1}] & \\ \mathfrak{m}_{n+1}[j] & \longrightarrow\!\!\!\longrightarrow & \mathfrak{m}'_n[j] & \longrightarrow\!\!\!\longrightarrow & \mathfrak{m}^\sim_n[j] & \longrightarrow\!\!\!\longrightarrow\!\!\!\longrightarrow\!\!\!\longrightarrow & \mathfrak{m}_n[j]. \end{array}$$

On the right hand side one reads the maximum over $0 \le u < m[n]$. To execute the last step we shall need lemma 15. For the sake of completeness we will list the steps that that are necessary to verify the hypothesis of lemma 15. Suppose

$\mathfrak{w}_n[j] \le \mathfrak{M}\eta(n)^{1-\gamma/(l-4-\beta)}$, for some constant $\mathfrak{M} > 0$ which shall be determined below. Analoguously to proposition 16 we obtain under the action of h^{-1} the following estimate

(5-7) $$\mathfrak{w}^{\wedge}_n[j] \le \mathfrak{z}^{1/4}(\mathfrak{w}_{n+1}[j] + \|\partial^2 \upsilon[0]\| + \ldots + \|\partial^2 \upsilon[l']\|)$$

provided $\mathfrak{w}_{n+1}[j] \le \mathfrak{z}_0$. Similarly for the second step one finds

(5-8) $$\mathfrak{w}^{\sim}_n[j] \le \mathfrak{w}^{\wedge}_n[j]\Gamma_{x,n}{}^{j-1}.$$

It is clear that $\mathfrak{w}_i[j] \le \mathfrak{M}\eta(n)^{1-\gamma/(l-4-\beta)}$, for all $j = 1,\ldots,\psi$, $i \in \mathbb{N}$ (and $\mathfrak{M}$ big enough). Thus we are lead to

$$\mathfrak{w}^{\sim}_n[j] \le \Gamma_{x,n}{}^{j-1}\mathfrak{M}\mathfrak{z}^{1/4}\eta(n+1)^{1-\gamma/(l-4-\beta)} + (l+2)\eta_{j+1}(n).$$

This, $\eta(n+1) \le \eta(n)\Delta_n{}^{l-4-\beta}$ and $\Delta_n{}^{l-j-3-\beta-\gamma} \le 1$ leads to

$$\mathfrak{w}^{\sim}_n[j] \le \mathfrak{M}\mathfrak{D}_n\mathfrak{z}^{1/4}\Delta_n\eta(n)^{1-\gamma/(l-4-\beta)} + (l+2)\eta_{j+1}(n),$$

for constants $\mathfrak{D}_n$ which in the following (as in the previous proposition) will take account of errors which are dominated by an expression of the form $1 + \mathfrak{d}m[n]^2 m[n-1]\eta_2$. The last inequality will be multiplied with mc. Hence

$$mc\mathfrak{w}^{\sim}_n[j] \le \mathfrak{T}^*\mathfrak{M}\mathfrak{D}_n\mathfrak{z}^{-1/4}\eta(n)^{1-2\gamma/(l-4-\beta)} + \mathfrak{z}^{\sim}_0/2,$$

if Θ^* is small enough (as it was given in proposition 16), where we used $\Delta_n m \le \mathfrak{z}^{-1/2}$ (for almost all $n \in \mathbb{N}$) and $c\eta(n)^{\gamma/(l-4-\beta)} \le \mathfrak{T}^*$. By shrinking Ξ we can achieve that $\mathfrak{T}^*\mathfrak{M}\mathfrak{D}_n\mathfrak{z}^{-1/4}\eta(n)^{1-2\gamma/(l-4-\beta)} \le \mathfrak{z}^{\sim}_0/2$. Thus $\mathfrak{w}^{\sim}_n[j]$ satisfies the hypothesis of lemma 15: $mc\mathfrak{w}^{\sim}_n[j] \le \mathfrak{z}^{\sim}_0$.

Back in proposition 16 we had $y^*[u](x) = -\nu_{10}[u](x,x')$ (dropping the index n), from which one finds with the implicit function theorem $x' = g'[u](x)$, for functions $g'[u]$: $J^{\sim} \to J^*[u]$. The derivatives of $g'[u]$ are estimated in the same way as it was done for $g[u]$. Thus we are led to apply lemma 15. The conditions of lemma 15 are satisfied as we have seen in proposition 16. We obtain

$$\|\partial_x g'[u]\| \le 1 + \theta' u^2 c\eta + uc\mathfrak{z}^{\sim} + \mathfrak{P}(\theta' u^2 c\eta + uc\mathfrak{z}^{\sim}),$$
$$\|\partial_x{}^r g'[u]\| \le \theta' u^2 c\eta + uc\mathfrak{z}^{\sim} + \mathfrak{P}(\theta' u^2 c\eta + uc\mathfrak{z}^{\sim}),$$

for $r = 2,\ldots,\psi$ and some polynomial $\mathfrak{P}$ that begins with quadratic terms. It follow for the derivatives of the composition $h^{\sim} \circ g'[u]$:

$$\partial^r(\mathfrak{h}^{\sim} \circ g'[u]) = \textstyle\sum_{1 \le t \le r} ((\partial^t \mathfrak{h}^{\sim}) \circ g'[u])\mathfrak{P}_t[u],$$

where $\mathfrak{P}_t[u]$ is a sum of products $\prod_{1\le s\le t} \partial^{i[s]}g'[u]$, with $i[s] \ge 1$ and $\sum_{1\le s\le t} i[s] = r$. In particular, $\mathfrak{P}_t[u]$ containes higher than first order derivatives if $t < r$. We therefore get

$$\|\partial^r(h^{\sim}\circ g'[u])\|_{J^*[u]} \le \mathfrak{w}^{\sim}_n[r] + \sum_{1\le t\le r} \mathfrak{w}^{\sim}_n[t]\big(\theta' u^2 c\eta + uc\mathfrak{y}^{\sim} + \mathfrak{B}(\theta' u^2 c\eta + uc\mathfrak{y}^{\sim})\big)$$

and therefore

$$\mathfrak{w}_n[j] \le (\mathfrak{W}\mathfrak{D}_n\mathfrak{z}^{1/4}\Delta_n\eta(n)^{1-\gamma/(1-4-\beta)} + \eta(n))(1 + \mathfrak{W}^{\wedge} m[n]\eta(n)),$$

for some constant $\mathfrak{W}^{\wedge} > 0$. We may assume that the constants $\mathfrak{D}_n = 1 + \mathfrak{b} m[n]^2 m[n-1]\eta(n)$ are less than $\mathfrak{z}^{1/8}$, say (if Θ^* (and hence Θ) is small enough). Thus $\mathfrak{D}_n\mathfrak{z}^{1/4}\Delta_n \le \mathfrak{z}^{-1/8}$, since $\Delta_n \le \mathfrak{z}^{-1/2}$ for almost all $n \in \mathbb{N}$ (the legitimacy of this replacement was discussed in lemma 10). We estimate $m[n]\eta(n)$ roughly by Θ^* and set

$$\mathfrak{W} = ((1 + \mathfrak{W}^{\wedge}\Theta^*)^{-1} - \mathfrak{z}^{-1/8})^{-1}.$$

Then obviously the induction hypothesis $\mathfrak{w}_n[j] \le \mathfrak{W}\eta(n)^{1-\gamma/(1-4-\beta)}$ is satisfied. The $\mathfrak{h}_0(i)$ map μ onto intervals $I_0(i)$ whose length are uniformly bounded from above and below. By an affine streching $I_0(i)$ are normalised to the unit interval. Denote by $\mathfrak{h}(i)$ the new functions. We have therefore proven that the sequence $\mathfrak{h}(i)$ is contained in a compact subset in C^{ψ} and thus by the theorem of Arzela - Ascoli converging uniformly in $C^{\psi-1}$ (if necessary by passing to a subsequence) to a limit function $\mathfrak{h}(x)$ whose $(\psi-1)$-derivative is Lipshitz. It follows from the construction that $\mathfrak{H} = \mathfrak{h}\circ\pi_x$ is the transformation which realises the conjugacy $\mathfrak{H}\circ\mathfrak{K}\circ\mathfrak{H}^{-1} = R_\omega$. □

6. Proof of the theorems.

Proof of theorem 3. We combine proposition 7 and the lemmas 9 - 17. The size of the neigbourhood of the simple line depends upon the Liouville exponent γ and the Liouville constant C. All that remains to do, is to check that condition (i) of proposition 7 agrees with the bound on the Liouville exponent. By lemma 12 $\Delta_{n+1}{}^{-1} \le \mathfrak{C}^{\wedge} M_{n+1}{}^{\gamma}$; for a new constant $\mathfrak{C}^{\wedge} \sim \mathfrak{C}$ (to level out the effect caused through replacing $m[n]$ by $\Delta_n{}^{-1}$)and $M_{n+1} = \Delta_n{}^{-1}M_n$. Therefore

$$\Delta_{n+1}{}^{-2}\Delta_n{}^{-1}\eta(0)M_{n+1}{}^{4+\beta-1} \le \Xi\mathfrak{C}^{\wedge 2}M_{n+1}{}^{2\gamma+4+\beta-1}\Delta_n{}^{-1} = \Xi\mathfrak{C}^{\wedge 2}\Delta_n{}^{-(2\gamma+5+\beta-1)}M_n{}^{2\gamma+4+\beta-1}$$

$$\le \Xi\mathfrak{C}^{\wedge 7+2\gamma+\beta-1}M_n{}^{\gamma(2\gamma+5+\beta-1)}M_n{}^{2\gamma+4+\beta-1}.$$

We see that poposition 7 (i) is satisfies if $\Xi \le 2^{-1}\Theta^{l-7-2\gamma-\beta}$ ($\Theta/2$ instead of Θ because we used Δ's instead of m's) and if

$$2\gamma^2-\gamma(l-7-\beta)-(l-4-\beta) \le 0$$

which holds true exactly if $\gamma \le \gamma^*(l) = (l-7-\beta)/4 + \sqrt{[(l-3-\beta)^2/16 + 1/2]}$ □

Proof of theorem 1. Let Φ be a lift of φ and set

$$(U^*,T^*) = (\Phi^{q[0]}R^{p[0]}, \Phi^{q[1]}R^{p[1]}) = (\Phi, \Phi^{m[0]}R).$$

Clearly this is a pair of commuting twist maps and furthermore we can be normalise them. This can be done fo instance by an affine coordinate transform Λ^* whose Jacobian has only in the diagonal non zero entries. Thus we set

$$(U_0,T_0) = (\Lambda^*U^*\Lambda^{*-1},\Lambda^*T^*\Lambda^{*-1}).$$

By theorem 2 there exists an $[L-3-\beta-\gamma]$ - differentiable curve that is invarint under (U_0,T_0) and which has rotation number $\omega' = [m[1],m[2],...]$. And therefore Φ has an invariant curve in the class of smoothness with rotation number $\omega = [m[0],[1],...])$. This concludes the proof of the theorem. □

References.

[1] Arnold, V. I.; Proof of A. N. Kolmogorov's theorem on the preservation of quasiperiodic motions under small perturbations of the Hamiltonian. Usp. Mat. Nauk SSSR 18, no 5, pp. 13 - 40, (1963).

[2] Hawkins - Schmidt; On C^2-diffeomorphisms on the circle which are of type III_1. Invent. math. 66., pp. 511 - 518, (1982).

[3] Herman, M. R.; **Sur les courbes invariantes par les diffeomorphisms de l'anneau ....** Asterisque 103 - 104, (1983).

[4] MacKay, R. S.; A renormalisation approach to invariant circles in area preserving maps. Physica 7 D, pp. 283 - 300, (1983).

[5] Mather, J. N.; Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. Topology 21, pp. 457 - 467, (1982).

[6] Moser, J. K.; On invariant curves of area preserving mappings of an annulus. Akad. der Wiss. Göttingen, pp. 1 - 20, (1962).

[7] Rand, D. A.; Universality for critical golden circle maps and the breakdown of dissipative golden invariant tori. Warwick preprint 1984.

[8] Rüssmann, H.; On the existence of invariant curves of twist mappings of an annulus. Springer Lecture Notes # 1007, pp. 677 - 718, (1983).

On Gibbs' and equilibrium states.

N. T. A. Haydn

Revised, February 1986.

Mathematics Institute, University of Warwick, Coventry CV4 7AL.

Abstract: Ruelle and Capocaccia gave a new definition of Gibbs states on Smale spaces. Equilibrium states of suitable functions there on are known to be Gibbs states. The converse is discussed in this paper, where the problem is reduced to shift spaces and there solved by constructing suitable conjugating homeomorphisms in order to verify the conditions for Gibbs states which Bowen gave for shift spaces, where the equivalence to equilibrium states is known.

Let (Ω,T) be a Smale space, then for any continous function $F\colon \Omega \to \mathbb{R}$ the pressure $P(T,F)$ can be defined by the variational principle

$$P(T,F) = \sup_\rho(h_T(\rho) + {}_\Omega\!\int F\, d\rho),$$

where ρ runs over all T-invariant probability measures over Ω. Here $h_T(\rho)$ is the measure theoretic entropy with respect to T and ρ. A measure, for which the supremum is attained, is called an equilibrium state. For every Hölder continuous F there is a unique equilibrium state if (Ω,T) is topologically mixing (cf. [1] Theorem 1.22).

Let $d(\cdot,\cdot) = d_\Omega(\cdot,\cdot)$ denote a metric on Ω. A map ψ from some open $U \subset \Omega$ into Ω is called *conjugating*, if $d(T^k \circ \psi(x), T^k(x)) \to 0$ for $|k| \to \infty$ uniformly in $x \in U$. Actually, as one may derive from Lemma 4, the distance decreases in a uniformly exponential way with a properly chosen metric.

Definition 1: (cf. [2] and [4].) *Let F be a Hölder continuous real valueo function on Ω. A probability measure ν is called a **Gibbs state** for F if*

$$\int_U \tau\circ\psi \, g \, d\nu = \int_{\psi(U)} \tau \, d\nu,$$

where

$$g(z) = \exp \sum_{k\in\mathbf{Z}} (F\circ T^k\circ\psi(z) - F\circ T^k(z)),$$

for all bounded and measurable functions $\tau: \psi(U) \to \mathbf{R}$ *and all conjugating homeomorphisms* $\psi: U \to \psi(U)$, *where* $U = U_\psi$ *is an open set in* Ω.

Equilibrium states for F are also Gibbs states for F. This is proven in Ruelle's book [4], theorem 7,18. The converse, there referred to as an open question (cf. [1] p. 170), will be demonstrated in this paper, i.e.

Theorem 2: *Let* (Ω,T) *be a Smale space and* $F \in C^\theta(\Omega)$, *for* $\theta \in (0,1)$, *i. e.,* F *is a Hölder continuous real valued function over* Ω *with exponent* θ, *and let* ν *be a Gibbs state for* F. *Then there is a number* $t \in \mathbf{N}$ *such that* ν *is invariant under* T^t *and is an equilibrium state for* $F_t = \sum_{0\le r<t} F\circ T^r$ *on* (Ω,T^t).

The proof is by a sequence of lemmas. We begin introducing a Markov partition on the Smale space (Ω,T), which gives rise to a shift space (Σ_A,σ). Most of the proof will be treated on this symbolic level. In order to see that every Gibbs measue on (Ω,T) lifts in a well-defined way to one on (Σ_A,σ) it is shown that the boundary set of any Markov partition has measure zero. This is in Lemma 5 first done for T-invariant Gibbs states and then in Proposition 6 generalised for non-invariant Gibbs measures. For constructing conjugating homeomorphisms on (Σ_A,σ) an obvious method exists, which will be used at three stages in this paper: first in Lemma 7 to estimate the measure of cylinder sets; secondly in Lemma 9 to show that the wandering set has measure zero, and finally in Lemma 10 to prove that Gibbs states on shift spaces (Σ_A,σ) with σ acting topologically mixing, are invariant under σ. This together (Proposition 11), provided σ is topologically mixing, verifies the definition Bowen gave for Gibbs states (cf.

[1], Theorem 1.2), for which it is known that they are equilibrium states on (Σ_A,σ). By standard results the same holds true for (Ω,T). To justify the construction of conjugating homeomorphisms on (Σ_A,σ) we begin in Lemma 4 by proving that they can be 'pushed down' to conjugating homeomorphisms on (Ω,T).

From now on ν denotes a Gibbs state for F on (Ω,T).

Let x be any point in Ω, then there exist a positive number δ and a number $\lambda \in (0,1)$, such that

$$d(T^j(y),T^j(z)) \leq \ell\cdot\lambda^{|j|}d(y,z) \quad \text{if} \quad \begin{cases} y,z \in V^+_x(\delta) & \text{for } j \geq 0, \\ y,z \in V^-_x(\delta) & \text{for } j < 0, \end{cases}$$

for a positive constant ℓ and where $V^+_x(\delta)$ $(V^-_x(\delta))$ denotes the stable (unstable) manifold through x cut off at distance δ. The constant ℓ can be assumed to be 1, $d(\cdot,\cdot)$ is then an adapted metric, which always exists. Furthermore T is expansive i.e. there exists a positive constant ε, such that for two points $x,x' \in \Omega$, with $x \neq x'$ there is an index $n \in \mathbf{Z}$ for which $d(T^n(x),T^n(x')) > \varepsilon$. There exist Markov partitions with arbitrary small diameter (cf. [1]). We choose one, $\{R[j]:\ j \in \mathcal{A}\}$ over an alphabet $\mathcal{A}$, such that $\operatorname{diam} R[j] < \varepsilon/2$ for all $j \in \mathcal{A}$. The sets $R[j]$ are called rectangles. Let A be the corresponding transition matrix and define the shift space

$$\Sigma_A = \{z:\ z = \{z_i : \ i \in \mathbf{Z}\},\ A[z_i,z_{i+1}] = 1 \ \text{ for all } \ i \in \mathbf{Z}\},$$

which, endowed with the metric $d_\Sigma(x,y) = \lambda^k$, where $k = \max\{j:\ x_i = y_i$ for all $|i| < j\}$, is a metric space. The surjection $\pi:\ \Sigma_A \to \Omega$ is Lipshitz continuous with constant, say, L. See also [4] p. 130. The two-sided shift σ on Σ_A is defind by $\sigma(z) = z'$ where $z'_i = z_{i+1}$ for all $i \in \mathbf{Z}$, and covers T on Ω: $T\circ\pi = \pi\circ\sigma$. We will need

Lemma 3. *Let* $\varepsilon_0, \varepsilon_1$ *be expansive constants, then*

(i) Given $\delta > 0$ *and let* M_δ *be the maximal number, such that* $y,y' \in \Omega$

with $d_\Omega(y,y') < \delta$ *implies* $d_\Omega(T^j(y),T^j(y')) < \varepsilon_0$ *for all* $|j| < M_\delta$, *then* $M_\delta \to \infty$ *as* δ *approaches zero.*

(ii) Given $M \in \mathbb{N}$ *and let* δ'_M *be the smallest number, such that* $y,y' \in \Omega$ *with* $d_\Omega(T^j(y),T^j(y')) < \varepsilon_1$ *for all* $|j| < M$ *implies* $d_\Omega(y,y') < \delta'_M$, *then* $\delta'_M \to 0$ *as* $M \to \infty$.

The **proof** is easy; see e.g. [1].

Lemma 4. *Let* φ: $U \to \varphi(U)$, *for* $U \subset \Sigma_A$ *be a uniformly continuous conjugating homeomorphism for* σ, *then there are a finite decomposition of* $U = \bigcup_i U_i$ *and conjugating homeomorphisms* ψ_i: $V_i = \pi(U_i) \to \psi_i(V_i)$, $V_i \subset \Omega$, *for* T, *which are projections of* φ *on* U_i, *i.e.* $\psi_i \circ \pi = \pi \circ \varphi$ *on* U_i *for all* i.

Proof. There is a $n^* \in \mathbb{N}$, such that $(\varphi(z))_i = z_i$ for all $|i| \geq n^*$ and $z \in U$. Otherwise there would be a sequence $\{z^k\colon k \in \mathbb{N}\}$, such that $(\varphi(z^k))_{i[k]} \neq z^k{}_{i[k]}$ for a sequence $|i[k]| \to \infty$ as $k \to \infty$. Or, $d_\Sigma(\sigma^{i[k]} \circ \varphi(z^k), \sigma^{i[k]}(z^k)) = 1$ for all $k \in \mathbb{N}$, but this contradicts the fact that $d_\Sigma(\sigma^j \circ \varphi(z), \sigma^j(z))$ converges to zero uniformly in z.

By uniform continuity there exists a $\delta \in (0,1)$, such that $d_\Sigma(x,y) \leq \delta$ implies $d_\Sigma(\varphi(x), \varphi(y)) \leq \lambda^{n^*}$. Now set

$$n = \max(n^*, [\log \delta / \log \lambda + 1]).$$

If for two points $x,y \in U$ we have $x_i = y_i$ for $|i| < n$ the same is true for their images under φ: $(\varphi(x))_i = (\varphi(y))_i$ for all $|i| < n$. Select a finite number of points $\{y^i\colon i \in \mathcal{J}\} \subset U$, with $|\mathcal{J}| < \infty$, and set $U_i = \{x\colon x_j = y^i{}_j \ \forall \ |j| < n\} \cap U$. A good choice of this set yields $U = \bigcup_{i \in \mathcal{J}} U_i$.

The rectangles $R[i], R[j]$ are said to be related if $R[i] \cap R[j] \neq \emptyset$. Denote by $\partial^+ R[j]$ and $\partial^- R[j]$ the forward and backward oriented parts of the boundary of $R[j]$, and set $\partial R = \bigcup_{j \in A} (\partial^+ R[j] \cup \partial^- R[j])$. Call the entire boundary set $K = \bigcup_{k \in \mathbb{Z}} T^k(\partial R)$.

The components of two points in $\pi^{-1}(K) \subset \Sigma_A$ always denote pairs of related rectangles. The same is true for their images under φ if we restrict to U_i for all $i \in J$. To see this choose any two points $x, x' \in U_i$ for which $\pi(x) = \pi(x')$, then $R[x_j]$ and $R[x'_j]$ are related for all $j \in \mathbb{Z}$. By construction of U_i their images have the same components for $|j| < n$ and so $R[(\varphi(x))_j]$ and $R[(\varphi(x'))_j]$ are again related for all $j \in \mathbb{Z}$. Since

$$d_\Omega(T^j \circ \pi \circ \varphi(x),\ T^j \circ \pi \circ \varphi(x')) \le \operatorname{diam}(R[(\varphi(x))_j] \cup R[(\varphi(x'))_j]) < \varepsilon,$$

for all $j \in \mathbb{Z}$, we conclude, that $\pi \circ \varphi(x) = \pi \circ \varphi(x')$ since ε is an expansive constant (cf. [3]).

Define on $V_i = \pi(U_i)$ ψ_i: $x \to \pi \circ \varphi(\pi^{-1}(x) \cap U_i)$. The sets V_i are not necessarily open, but contain open subsets. The maps ψ_i will turn out to be conjugating and injective maps on V_i. Set $\beta = [1 + |\log(\varepsilon/(4L))/\log \lambda|]$ and take for each i a finite cover $\{U_{i,k}: k \in \mathcal{K}_i\}$ of U_i with $|\mathcal{K}_i| < \infty$, where $U_{i,k} = \{x: x_j = z^k{}_j \ \forall \ |j| < n+\beta\}$ for finite subsets $\{z^k: k \in \mathcal{K}_i\}$ chosen as before. It remains to be shown that the maps ψ_i are continuous. This is done for each $U_{i,k}$, the regionbeing extended to the closures of the $\pi(U_{i,k})$ in $\pi(U_i)$.

Choose any two points $x, x' \in U_{i,k}$ with $y = \pi(x)$ and $y' = \pi(x')$, i.e. $y, y' \in V_{i,k} = (\mathrm{cl.}\ \pi(U_{i,k})) \cap \pi(U_i)$. Now we apply Lemma 3 (i) setting $\varepsilon_0 = \varepsilon/2$ (and restrict to δ, so that $M_\delta > n+\beta$). From the construction of the $V_{i,k}$ it is clear that $T^j \circ \pi \circ \varphi(x)$ and $T^j \circ \pi \circ \varphi(x')$ travel under T through the same rectangles for $|j| < n+\beta$. Estimate

$$d_\Omega(T^j \circ \psi_i \circ \pi(x),\ T^j \circ \psi_i \circ \pi(x'))$$

$$\le L\{d_\Sigma(\sigma^j \circ \varphi(x), \sigma^j(x)) + d_\Sigma(\sigma^j(x'), \sigma^j \circ \varphi(x'))\} + d_\Omega(T^j \circ \pi(x), T^j \circ \pi(x'))$$

$$\le 2L\lambda^{|j|-n} + \varepsilon/2 \ \le \varepsilon$$

if $n+\beta \le |j| < M_\delta$. Hence $d_\Omega(T^j \circ \psi_i(y), T^j \circ \psi_i(y')) \le \varepsilon$ for all $|j| < M_\delta$ and we conclude using Lemma 3 (ii) with $\varepsilon_1 = \varepsilon$, that $d_\Omega(\psi_i(y), \psi_i(y')) \le \delta'$ where

$\delta' \to 0$ as δ approaches zero. Moreover, ψ_i is a continuous homeomorphism onto the boundaries $\partial(\pi(U_{i,k})) \cap \partial(\pi(U_{i,l})) \cap \pi(U_i)$ for all k,l. This gives us the maps ψ_i defined on $\pi(U_i)$ in the desired form. Obviously they are conjugating.

□

Lemma 5. *If* ν^* *is a* T-*invariant Gibbs measure on* Ω , *then* $\nu^*(K) = 0$, *where* $K = \bigcup_{k\in\mathbf{Z}} T^k(\partial R)$.

Proof. Suppose $\nu^*(K) > 0$, then so is $\nu^*(\partial R) > 0$ and thus either $\nu^*(\bigcap_{i\in\mathbf{N}} T^i(\partial^+ R)) > 0$ or $\nu^*(\bigcap_{i\in\mathbf{N}} T^{-i}(\partial^- R)) > 0$. Assume the first case holds, and let $K^* = \bigcap_{i\in\mathbf{N}} T^i(\partial^+ R)$. Denote by $B_\zeta(z)$ the ball around z with radius ζ. The set K^* is invariant under T and is compact; hence there is a point $z \in K^*$, such that $\nu^*(B_\zeta(z) \cap K^*) > 0$ for all $\zeta > 0$. The points conjugated to z are dense in Ω and for each of them there exists a conjugating homeomorphism defined in a neighbourhood of z (cf. [2]). So there is a $w \in \Omega$ and a $\delta > 0$, so that $B_{2\delta}(w) \cap K^* = \emptyset$, and a conjugating homeomorphism ψ defined on $B_\zeta(z)$ for a positive ζ, such that $\psi(z) \in B_{\delta/2}(w)$ and $\psi(B_\zeta(z)) \subset B_\delta(w)$. Set $D = \psi(B_\zeta(z) \cap K^*)$, then we have $\nu^*(D) > 0$, because ν^* is a Gibbs state. Since ψ is conjugating, there exists a $n' \in \mathbf{N}$ such that $d(T^i \circ \psi(y), T^i(y)) < \delta$ for all $y \in B_\zeta(z)$ and $|i| \ge n'$. So $\mathrm{dist}(T^i(D), K^*) < \delta$ for K^* is invariant under T. But by construction $\mathrm{dist}(D, K^*) > \delta$, hence $T^i(D) \cap D = \emptyset$ for all $i \ge n'$. The collection $\{T^{in'}(D): i \in \mathbf{N}\}$ consists of pairwise disjoint sets, which have, since ν^* is T-invariant, the same, positive measure. Hence, the measure of their union diverges, which contradicts the normalisation $\nu^*(\Omega) = 1$.

□

Proposition 6. *For a Gibbs measure* ν *on* (Ω, T) *it holds* $\nu(K) = 0$.

Proof. Assume $\nu(K) > 0$, then there must be an integer j so that

$\nu(T^j(\partial R)) > 0$ and thus either $\nu(T^j(\partial^+R)) > 0$ or $\nu(T^j(\partial^-R)) > 0$. Suppose we have the situation $j = 0$ and $\nu(\partial^-R) > 0$. Then define a sequence of new measures

$$\nu_n = n^{-1}\cdot\sum_{0\le i<n} \nu\circ T^i.$$

Since ν is Gibbs it is clear that $\nu\circ T^i$ is Gibbs as well for all $i \in \mathbf{Z}$. By the convexity property (cf. [2]) all ν_n, for $n \in \mathbf{N}$, are Gibbs measures. Let ν^* be a limit point of $\{\nu_n\colon n \in \mathbf{N}\}$ and $n[j]$ be a subsequence in $\mathbf{N}$ so that $\nu_{n[j]}$ converges to ν^*. For the backward oriented boundary $\partial^-R \subset T(\partial^-R)$ holds, and therefore $\nu(\partial^-R) \le \nu_{n[j]}(\partial^-R)$. We may treat ∂^-R as a compact set and have therefore $\limsup_j \nu_{n[j]}(\partial^-R) \le \nu^*(\partial^-R)$, which is $\le \nu^*(K) = 0$ since $\partial^-R \subset K$.

□

The function F, acting on Ω, induces a Hölder continuous real valued function $f = F\circ\pi$ on Σ_A, which is exponentially decreasing with $\alpha = \lambda^\theta \in (0,1)$. Let

$$\mathrm{var}_k\, f = \sup\{|f(x) - f(y)|\colon\ x,y \in \Sigma_A \text{ such that } x_i = y_i\ \forall\ |i| < k\}$$

and set

$$\|f\| = \max(\|f\|_\infty, \sup_{k\in\mathbf{Z}} \alpha^{-|k|}\mathrm{var}_k\, f).$$

Restricted to $\Omega\backslash K$, the map π^{-1} is one to one, that is, by Proposition 6, π^{-1} is defined ν-almost everywhere. Define the measure μ on Σ_A by $\mu(V) = 0$ if $V \subset \pi^{-1}(K)$, and $\mu(V) = \mu(V \cap \pi^{-1}(\Omega\backslash K)) = \nu(\pi(V))$ for all other $V \subset \Sigma_A$. Then μ is a Gibbs state for f, since ν is one for F and all bounded and measurable (test)functions τ on Σ_A can be written as $\tau = \tau^\sim\circ\pi$ almost everywhere with $\tau^\sim$ bounded and measurable functions on Ω.

The cylinder in Σ_A determined by the string $x_a, \dots, x_b$ will be denoted by $U(x_a, \dots, x_b)$ and for convenience we will write $\mu(x_a, \dots, x_b) = \mu(U(x_a, \dots, x_b))$.

Lemma 7: *Let* f *be a function on* Σ_A *which decreases exponentially fast* *with* $\alpha \in (0,1)$ *and let* σ *be topologically mixing. Then there is a constant* $C^* \in (0,\infty)$ *so that*

(1) $$\mu(x_1, \ldots, x_m) \cdot \exp(mP - \sum_{k\in[1,m]} f \circ \sigma^k(x)) \in [e^{-C^*}, e^{+C^*}]$$

for all $x \in \Sigma_A$ *and* $m \in \mathbf{N}$. *The real number* P *is called the pressure of* f.

Proof. To verify (1) we will construct a sequence of sets of conjugating homeomorphisms on Σ_A. Since σ is supposed to act topologically mixing, we have $A^n > 0$ if n is large enough. Let N be the smallest such integer and let

$$\mathcal{T}^*[m] = \{(a_1,\ldots,a_m): \; a_i \in \mathcal{A}, \; A[a_i,a_{i+1}] = 1 \;\; \forall \; i \in [1, m)\}$$

for the set of all m-strings. The pressure of f is $P = \lim_{m\to\infty} P_m$, where

$$P_m = m^{-1}\log Z_m$$

and

$$Z_m = \sum_{a\in\mathcal{T}^*[m]} \exp \sup\{\sum_{k\in[1,m]} f\circ\sigma^k(z): \; z_i = a_i \;\; \forall \; i \in [1,m]\}.$$

Let us first summon a technical lemma.

Lemma 8: *There exists a number* b, *such that* $|P_m - P| \leq b/m$ *for all* $m \geq 1$.

Proof. Set

$$Z^\sim_m = \sum_{z\in\mathcal{P}[m]} \exp \sum_{k\in[1,m]} f\circ\sigma^k(z),$$

where $\mathcal{P}[m] = \{x: \sigma^m(x) = x, \; x \in \Sigma_A\}$ is the set of all m-periodic points. Then, using the maximum matrix norm $\|A\| = \max_{i,j} |A[i,j]|$, it follows

$$Z^\sim_{m+N} \leq Z_m \|A^N\|^2 \exp(2N\|f\|_\infty).$$

Furthermore

$$Z_m \leq Z^\sim_{m+N} \exp(2\|f\|/(1-\alpha)).$$

Let $c_1 = 2\|f\|(N + 1/(1-\alpha)) + 2\cdot\log\|A^N\|$, then these two inequalities combine to $|\log(Z^\sim_{m+N} Z_m^{-1})| \leq c_1$. As can easily be seen, the pressure may be defined by $\lim_m m^{-1}\log Z_m$ and as well by $\lim_m m^{-1}\log Z^\sim_m$, which is the same. From [4] corollary 7.25 we know that $|\log Z^\sim_m - mP| \leq c_2 t^m$ for constants $c_2 > 0$

and $t \in (0,1)$, that is $|\log Z^{\sim}_{m+N} - mP| \le c_2 + N \cdot P$. Now set $b = N \cdot P + c_1 + c_2$.

□

Recall that $A^N > 0$ since σ is assumed to be topologically mixing. Choose any $x \in \Sigma_A$ and $m \in \mathbb{N}$ and fix them throughout the rest of the proof. Now we construct a collection of conjugating maps, which are all uniformly continuous and which depend on x and m. Set $U = \{z: \; z_i = x_i \; \forall \; i \in [1,m]\}$ and choose any string $\omega \in \mathcal{T}^*[m]$. Now we define a conjugating map φ depending on ω. Set for all $z \in U$

(i) $(\varphi(z))_i = \omega_i$ for all $i \in [1,m]$

(ii) $(\varphi(z))_i = z_i$ for all $i \in (-\infty, -N] \cup [m+N, +\infty)$

(iii) set $U_\omega = \{z: \; z_i = (\varphi(x))_i = \omega_i \; \forall \; i \in [1,m]\}$, and take a covering of U by $|A|^2$ (not necessarily non-empty) sets: $U(s,t) = U \cap \{z: \; z_{-N} = s, z_{m+N} = t\}$, for all $s,t \in A$. The sets $U_\omega(s,t)$ are defined analogously. For their measure we obtain

(2) $$\mu(U_\omega) = \sum_{s,t \in A} \mu(U_\omega(s,t)) = \sum_{s,t \in A} \mu(\bigcup_k U_\omega^k(s,t)),$$

where $U_\omega^k(s,t)$ are at most $\|A^N\|^2$ disjoint sets of points with the same symbols on the places in $[-N, m+N]$. Altogether they cover $U_\omega(s,t)$. Pick out the set $U_\omega^k(s,t)$ which realises the $\max_k \mu(U_\omega^k(s,t))$. This determines the components of $\varphi(\cdot)$ on the places with indices in the two intervals $(-N, 1)$ and $(m, m+N)$.

The map $\varphi: U(s,t) \to U_\omega^k(s,t)$ for all $s,t \in A$ is therefore completely defined. It is clear that φ is uniformly continuous on $U(s,t)$ for all $s,t \in A$, and conjugating, but it is not a homeomorphism. On $U(s,t)$ the map φ is finite but at most $(|A| \cdot \|A^N\|)^2$ to one and may therefore be decomposed in at most $(|A| \cdot \|A^N\|)^2$ homeomorphism. Moreover we derive from (2):

(3) $$\mu(\varphi(U)) \ge (|A| \cdot \|A^N\|)^{-2} \mu(U_\omega).$$

Finally we put

$S[m] = \{$all φ as constructed above with ω running over the whole $\mathcal{T}^*[m]\}$.

Let $\mathfrak{J} = \{1,2, \ldots, |S[m]|\}$ be a numbering of $S[m]$ and call the elements φ_j in it by an index. Set

$$g_j(x) = \exp \sum_{k \in \mathbf{Z}} (f \circ \sigma^k \circ \varphi_j(x) - f \circ \sigma^k(x))$$

and consider the following weighted sum:

$$\exp \sum_{k \in [1,m]} f \circ \sigma^k(x) \sum_{j \in \mathfrak{J}} \{g_j(x) \times \exp \sum_{k \in (-\infty,1) \cup (m,+\infty)} -(f \circ \sigma^k \circ \varphi_j(x) - f \circ \sigma^k(x))\}. \tag{4}$$

This is the same as

$$\sum_{j \in \mathfrak{J}} \exp \sum_{k \in [1,m]} f \circ \sigma^k \circ \varphi_j(x) = \sum_{j \in \mathfrak{J}} \exp(\sup\{\sum_{k \in [1,m]} f \circ \sigma^k(z) : z_i = (\varphi_j(x))_i \ \forall \ i \in [1,m]\} + r_j),$$

where the remainder r_j are estimated as $|r_j| \le \|f\|(1 + 2\alpha(1-\alpha^m)/(1-\alpha))$. The sum (4) lies therefore in the interval $Z_m[e^{-c}, e^c]$, where $c = \|f\|(1 + 2/(1-\alpha))$. Let $P_m = m^{-1} \log Z_m$ and set

$$\theta_m = \mu(x_1, \ldots, x_m) \cdot \exp(mP_m - \sum_{k \in [1,m]} f \circ \sigma^k(x)).$$

Observe that replacing P_m by P transforms θ_m into the expression to the left of (1). We use (4) to get rid of the factor mP_m in the exponential. Instead it appears a summation over j, expressing the sum over all m-strings in the definition of the pressure. So we end up with

$$\theta_m = \mu(x_1, \ldots, x_m) d_1 \sum_{j \in \mathfrak{J}} d_{2,j} g_j(x),$$

where $d_1 \in [e^{-c}, e^c]$, and

$$d_{2,j} = \exp \sum_{k \in (-\infty,1) \cup (m,+\infty)} (f \circ \sigma^k \circ \varphi_j(x) - f \circ \sigma^k(x)) \in [e^{-c'}, e^{c'}]$$

for all $j \in \mathfrak{J}$, with $c' = 2/(1-\alpha)$. For all $y \in U$ and $j \in \mathfrak{J}$ we estimate

$$\begin{aligned} &|\log(g_j(x) g_j(y)^{-1})| \\ &\le \sum_{k \in (-\infty,1) \cup (m,+\infty)} \{|f \circ \sigma^k \circ \varphi_j(x) - f \circ \sigma^k(x)| + |f \circ \sigma^k \circ \varphi_j(y) - f \circ \sigma^k(y)|\} \\ &\quad + \sum_{k \in [1,m]} \{|f \circ \sigma^k \circ \varphi_j(x) - f \circ \sigma^k \circ \varphi_j(y)| + |f \circ \sigma^k(x) - f \circ \sigma^k(y)|\} \\ &\le 2\|f\|(N + 2/(1-\alpha) + 1 + 2\alpha(1-\alpha^m)/(1-\alpha)). \end{aligned}$$

Set $c'' = 2\|f\|(N + 1 + 4/(1-\alpha))$ and it simplifies to

(5) $$|\log(g_j(x)g_j(y)^{-1})| \le c''.$$

Define the characteristic function of U

$$\tau_m^x(y) = \begin{cases} 1 & \text{if } y_i = x_i \text{ for all } i \in [1,m], \\ 0 & \text{otherwise.} \end{cases}$$

Then we have $U = \{y:\ y \in \Sigma_A,\ \ \tau_m^x(y) = 1\}$. For θ_m we may write

$$\theta_m = \sum_{j\in J} d_1 d_{2,j} g_j(x)_U\int \tau_m^x \, d\mu$$

$$\in \sum_{j\in J} {}_U\int g_j \tau_m^x \, d\mu [e^{-c-c'-c''}, e^{c+c'+c''}]$$

– the last step since inequality (5). To evaluate the integral on the right, we remember μ to be a Gibbs state for f. According to the note made when defining the conjugating maps, we decompose for every j the cylinder U into a finite number of sets $U[k,j]$, i.e $U = \bigcup_{k\in J[j]} U[k,j]$, where $\varphi_j|_{U[k,j]}$ are homeomorphisms and $J[j] \le (|A|\cdot\|A^N\|)^2$ for all j. Hence

$$\sum_{j\in J} {}_U\int g_j\cdot\tau_m^x \, d\mu = \sum_{j\in J} \sum_{k\in J[j]} {}_{U[k,j]}\int g_j\cdot\tau_m^x \, d\mu$$

$$= \sum_{j\in J} \sum_{k\in J[j]} \int \tau_m^x \circ \varphi_j^{-1} \, d\mu$$

where the integrals in the last line are taken over $\varphi_j(U[k,j])$. This expression lies in the interval $[(|A|\cdot\|A^N\|)^{-2}, (|A|\cdot\|A^N\|)^2]$. To see this, observe that the set $\{\omega[j]:\ \omega[j] = ((\varphi_j(x))_1, \dots, (\varphi_j(x))_m),\ j \in J\}$ is just $\mathcal{T}^*[m]$, that is $\Sigma_A = \bigcup_{j\in J} U_{\omega[j]}$. But $\bigcup_{j\in J} \varphi_j(U)$ does not cover the whole space Σ_A, since there was a choice in (iii) in defining φ_j. This and inequality (3) provides the lower bound. The upper can be explained by the possibility that the maps φ_j have to be decomposed in $|J[j]| \le (|A|\cdot\|A^N\|)^2$ homeomorphisms. It is therefore proven that $\theta_m \in [e^{-C}, e^{C}]$ with $C = c + c' + c'' + 2\cdot\log(\|A^N\|\cdot|A|)$ a constant independent of $x \in \Sigma_A$ and $m \in \mathbb{N}$. Finally we replace P_m by P, which is compensated for by increasing C to $C+b$, with the same b as in Lemma 8.

Set $C^* = C + b$ and (1) is proven for all $x \in \Sigma_A$ and $m \in \mathbb{N}$.

□

Lemma 9. *The wandering set has measure zero.*

Proof. The lemma is proven for the shift space and clearly holds then in the case of the underlying Smale space as well. Let A be the transition-matrix. It may be brought into the form

$$\begin{vmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,t} \\ 0 & A_{2,2} & \cdots & \vdots \\ 0 & \cdots\cdots & 0 & A_{t,t} \end{vmatrix},$$

where $A_{i,i}$ are irreducible $r_i x r_i$-matrices. The alphabet $\mathcal{A}$ splits into $(\mathcal{A}_1, \ldots, \mathcal{A}_t)$, with $|\mathcal{A}_i| = r_i$ for all i. The subshifts Σ_i generated by $A_{i,i}$ over $\mathcal{A}_i$ are topologically transitive and decompose into finitely many subshifts $\Sigma_{j,q}$ on which a power σ^s of the shift acts topologically mixing. We renumber the shifts $\Sigma_{i,q}$ to call them now Σ_i for $i \in [1,t']$ with $t' \geq t$. Let σ^s for convenience now be called σ and let N be the smallest number so that $A_{i,i}{}^N > 0$ for all $i \in [1,t']$ (with the new $A_{i,i}$) and denote by $P[i]$ the pressure of f on Σ_i. Points $x = (x_k)_{k \in \mathbb{Z}}$ in the wandering set of Σ_A are sequences with elements x_k not all in one subalphabet $\mathcal{A}_i$ but in several; however the indices $i[k]$ of $\mathcal{A}_{i[k]}$ to which the x_k belong never decrease as k increases. Suppose the wandering set has positive measure, then there exists a subset $U(\zeta) \subset \Sigma_A$, $\zeta = \zeta_a, \ldots, \zeta_b$ with positive measure and for which $U(\zeta) \cap U(\sigma^m(\zeta)) = \emptyset$ for all $m \in \mathbb{Z} \setminus \{0\}$, where $U(\zeta)$ denotes the cylinder in Σ_A which is determined by the string ζ.

Consider the cover $U(\zeta) = \bigcup U(\zeta_{a-1}, \zeta)$, where ζ_{a-1} runs over all symbols with $A[\zeta_{a-1}, \zeta_a] = 1$. If $\zeta_a \in \mathcal{A}_i$ and $\zeta_{a-1} \in \mathcal{A}_{i'}$ for $i < i'$ and if $\mu(\zeta_{a-1}, \zeta) > 0$ then we replace $\zeta = \zeta[0]$ by $\zeta[1] = \zeta_{a-1}, \zeta$ (again $\mu(\zeta_{a-1}, \zeta)$ stands for $\mu(U(\zeta_{a-1}, \zeta))$); otherwise we use for ζ_{a-1} any other symbol so that $\mu(\zeta_{a-1}, \zeta) > 0$. Repeat the same procedure for ζ_{b+1} and increase the index of the subalphabet if possible or leave it the same. In that way we get

$\zeta[2] = \zeta[1],\zeta_{b+1}$ with $\zeta_b \in \mathcal{A}_j$ and $\zeta_{b+1} \in \mathcal{A}_{j'}$, where $j \le j'$. Iterating this process one obtains a sequence of strings $\zeta[i]$ of increasing length and $\mu(\zeta[i]) > 0$ for $i \in \mathbb{N}\cup\{0\}$. At any stage the elements on both sides must remain in the same subalphabets, that is, there exists a number $M > \max(|a|, |b|)$ and i, j with $1 \le i < j \le t$, so that $\zeta_{-k} \in \mathcal{A}_i$ and $\zeta_k \in \mathcal{A}_j$ for all $k \ge M$. Set $\zeta_{-M},\dots,\zeta_M = \xi$ and

$$U^{\sim}(\xi) = \{x\colon\ x_{-k} \in \mathcal{A}_i,\ x_k \in \mathcal{A}_j,\ \text{for } k > M,$$
$$x_k = \xi_k \text{ for } |k| \le M \text{ and } A[x_k,x_{k+1}] = 1 \text{ for all } k \in \mathbf{Z}\}.$$

The same notation will be used whenever there are cylinders with elements on the negative side restricted to $\mathcal{A}_i$ respectively $\mathcal{A}_j$ on the positive side. By construction $\mu(U^{\sim}(\xi)) > 0$ and

$$\sigma^m(U^{\sim}(\xi)) \cap U^{\sim}(\xi) = \emptyset$$

holds true for all $m \in \mathbf{Z}\setminus\{0\}$. Suppose $P[i] \ge P[j]$. Fix $m > 2M+1+N$ and denote by $\{\gamma[p]\colon\ p \in [1,\mathfrak{p}]\}$, for a number $\mathfrak{p}$, the set of all m-strings in Σ_i. Let us construct conjugating homeomorphisms $\{\varphi_p\colon\ p \in [1,\mathfrak{p}]\}$. Select an m-string θ in Σ_j, then $\mu(U^{\sim}(\xi,\dots,\theta))$ is positive, for all choices of θ since μ is Gibbs. The dots denote room for $N-1$ symbols. Decompose

$$U^{\sim}(\xi,\dots,\theta) = \bigcup_{\eta,\eta'} U^{\sim}(\eta,\xi,\eta',\theta),$$

where η runs over all $(2N-1)$-strings in Σ_i and η' over all $(N-1)$-strings in Σ_j. Select a pair (η,η') so that

(6) $$\mu(U^{\sim}(\eta,\xi,\eta',\theta)) \ge c_0^{-1}\mu(U^{\sim}(\xi,\dots,\theta)),$$

with $c_0 = |\mathcal{A}|\cdot\|A^N\|\cdot\|A^{2N}\|$.

Define φ_p for $p \le \mathfrak{p}$ and set for $z \in U^{\sim}(\eta,\xi,\eta',\theta)$:

(i) $(\varphi_p(z))_k = z_k$ for $k \in (-\infty, -M-2N] \cup [M+N+m, \infty)$,

(ii) $(\varphi_p(z))_k = z_{k-m}$ for $k \in [-M+m, M+m]$,

(iii) $((\varphi_p(z))_{-M-N},\dots,(\varphi_p(z))_{-M-N+m}) = \gamma[p]$,

(iv) something fitting up in $(-M-2N, -M-N)$, $(-M-N+m, -M+m)$ and $(M+m, M+N+m)$, which will be specified in a moment.

Thus it follows for the Radon-Nykodym-derivative

$$\log d\varphi_p\mu/d\mu(z) = \sum_{k\in\mathbf{Z}} (f\circ\sigma^k\circ\varphi_p(z) - f\circ\sigma^k(z))$$

$$= \sum_{k\in(-\infty,-M-2N]\cup[M+N+m,\,\infty)} (f\circ\sigma^k\circ\varphi_p(z) - f\circ\sigma^k(z))$$

$$+ \sum_{k\in[-M,M]} (f\circ\sigma^{k+m}\circ\varphi_p(z) - f\circ\sigma^k(z)$$

$$+\sum_{k\in(-M-2N,-M-N)\cup(-M-N+m,-M+m)\cup(M+m,M+N+m)} f\circ\sigma^k\circ\varphi_p(z)$$

$$- \sum_{k\in(-M-2N,-M)\cup(M,M+N)} f\circ\sigma^k(z))$$

$$+ \sum_{k\in[-M-N,-M-N+m]} (f\circ\sigma^k\circ\varphi_p(z) - f\circ\sigma^{k+2(M+N)}(z)).$$

The first and second of these four sums are in modulus less than $4\|f\|/(1-\alpha)$, and the third and fourth together are less than $3\|f\|(N+4/(1-\alpha))$. Put $c = 3\|f\|(N+8/(1-\alpha))$. By construction we made the sets $\varphi_p(U^\sim(\eta,\xi,\eta',\theta))$ pairwise disjoint, and furthermore

$$\varphi_p(U^\sim(\eta,\xi,\eta',\theta)) \subset \sigma^m(U^\sim(\xi)),$$

that is, for different m we get disjoint sets. As pointed out in Lemma 7 part (iii), the fitting-up strings in (iv) may be chosen so that

$$\mu(\sigma^m(U^\sim(\xi))) \geq c_1^{-1}\mu(\textstyle\bigcup_{p\in[1,p]} \varphi_p(U^\sim(\eta,\xi,\eta',\theta))),$$

with $c_1 = |A|^2\cdot\|A^N\|^3$. So we have for any $z \in U^\sim(\eta,\xi,\eta',\theta)$:

$$\mu(\sigma^m(U^\sim(\xi))) \geq c_1^{-1}e^{-c}\mu(U^\sim(\eta,\xi,\eta',\theta))$$
$$\times \sum_{p\in[1,p]} \exp\sum_{k\in[1,m]} (f\circ\sigma^{k-M-N}\circ\varphi_p(z) - f\circ\sigma^{k+M+N}(z))$$

$$\geq c_1^{-1}e^{-c-P[j]}\mu(U^\sim(\eta,\xi,\eta',\theta))\exp\{P[j] - \sum_{k\in[1,m]} f\circ\sigma^{k+M+N}(z))\}$$
$$\times \sum_{p\in[1,p]} \exp\sum_{k\in[1,m]} f\circ\sigma^{k-M-N}\circ\varphi_p(z)$$

The same argument as in Lemma 7 provides a more general form of the inequalities (1), namely

$$\mu(U^{\sim}(\xi,\dots,\theta))\exp\{mP[j] - \sum_{k\in[1,m]} f\circ\sigma^{k+M+N}(z))\} \in \mu(U^{\sim}(\xi))[e^{-C^*}, e^{+C^*}],$$

for $z \in U^{\sim}(\xi,\dots,\theta)$. To see this, observe that the conjugating homeomorphisms there are to be constructed in exactly the same way. The result for Zeta-functions involving periodic points, which is cited in Lemma 8, works the same in this case too, since the number of periodic points with period h is the trace of the h-th power of the transitionmatrix, that is here $tr(A_{j,j}{}^h)$. To evaluate the pressure in that case we choose only points which are periodic in the components with index $\geq M+N$.

By following the argument after inequality (4) the summation over $p \in [1,p]$ gives rise to a factor $\exp(mP_m[i])$. Using (6) we end up with

$$\mu(\sigma^m(U^{\sim}(\xi))) \geq c_0{}^{-1}c_1{}^{-1}e^{-c-C^*}d_1\mu(U^{\sim}(\xi))\exp[m(P_m[i]-P[j])],$$

where d_1 is the same constant as in Lemma 7. Finally, once more Lemma 8, this transforms to

$$\mu(\sigma^m(U^{\sim}(\xi))) \geq c_0{}^{-1}c_1{}^{-1}e^{-c-C^*-b}d_1\mu(U^{\sim}(\xi))\exp[m(P[i]-P[j])]$$
$$\geq c_0{}^{-1}c_1{}^{-1}e^{-c-C^*-b}d_1\mu(U^{\sim}(\xi)),$$

since we supposed $P[i]-P[j]$ not to be negative. Summing up over m gives a contradiction to the normalisation condition $\mu(\Sigma_A) = 1$, hence $\mu(U^{\sim}(\xi)) = \mu(\zeta) = 0$, and therefore the lemma follows, since the wandering set of (Σ_A,σ) is contained in the one of (Σ_A,σ^s). In case $P[i] - P[j]$ happens to be negative we construct the conjugating homeomorphisms replacing m by $-m$ and get then the lower bound for $\mu(\sigma^m(U^{\sim}(\xi)))$ in the same way as described.

□

Lemma 10: *Let* μ *be a Gibbs state on* (Σ_A,σ), *then it is* σ*-invariant, if* σ *acts topologically mixing.*

Proof. Let $\mathcal{T}[b-a]$ be the set consisting of all possible strings $x_a,\dots,x_b$ of length $b-a$ for $a,b \in \mathbf{Z}$ and $a \leq b$.

Assume μ not to be σ-invariant, then there exists a set $B \subset \Sigma_A$ such that $|\mu(\sigma(B)) - \mu(B)| > 0$ and $\mu(B) > 0$. Suppose it is $\mu(\sigma(B)) \geq \rho\cdot\mu(B)$ for a number $\rho > 1$. By a covering argument we conclude that there must be a cylinder $U(\zeta)$ determined by the string $\zeta = x_a,\dots, \mathbf{x_0},\dots,x_b$ (bold characters denote the zero position), such that

$$\mu(\zeta^*) \geq \rho\cdot\mu(\zeta),$$

where $\zeta^* = x_a,\dots, \mathbf{x_1},\dots,x_b$ is the shifted string ζ. A partial covering of Σ_A will now be constructed, and then it will be shown that $\mu(\Sigma_A) \geq \sqrt{\sqrt{\rho}}$, which contradicts the normalisation of μ.

Let τ be an even number such that

$$16\|f\|\cdot\alpha^{\tau/2} \leq \log\rho,$$

and let us construct a cover of $U(\zeta)$ by cylinder sets of the form $U(\beta,\eta,\zeta,\eta',\beta')$, where $\beta,\beta' \in \mathcal{T}[\tau]$, $\eta,\eta' \in \mathcal{T}[N-1]$. There are strings β, β', η, η' so that

$$\mu(\beta,\eta,\zeta^*,\eta',\beta') \geq \rho\cdot\mu(\beta,\eta,\zeta,\eta',\beta').$$

One single * means the whole string is to be shifted. For the moment fix β, β', η, η', and proceed to construct a pair of conjugating homeomorphisms φ, φ' as follows. Choose any string $\omega \in \mathcal{T}[b-a]$ with $\omega = \omega_a,\dots, \boldsymbol{\omega_0},\dots,\omega_b$ and define for all $z \in U(\beta,\eta,\zeta,\eta',\beta')$

(i) $(\varphi(z))_i = \omega_i$ for all $i \in [a,b]$,

(ii) $(\varphi(z))_i = z_i$ for all $i \in (-\infty, a-N] \cup [b+N, \infty)$,

(iii) any strings $\theta,\theta' \in \mathcal{T}[N-1]$ to join up the ends in the intervals $(a-N,a)$ and $(b,b+N)$.

Secondly we define φ' on the cylinder $U(\beta,\eta,\zeta^*,\eta',\beta')$ and set for all z in it

(i') $(\varphi'(z))_i = (\omega^*)_i$ for all $i+1 \in [a,b]$,

(ii') $(\varphi'(z))_i = z_i$ for all $i+1 \in (-\infty, a-N] \cup [b+N, \infty)$,

(iii') the same strings θ and θ' as in (iii) to fill the two gaps.

We observe that $\varphi'\circ\sigma\circ\varphi^{-1} = \sigma$ on $U(\beta,\theta,\omega,\theta',\beta')$. Denote by $\chi(x_a,\dots, \mathbf{x_0},\dots,x_b)$ the characteristic function of $U(x_a,\dots, \mathbf{x_0},\dots,x_b)$. Since μ is by hypothesis a Gibbs state we conclude for any $x \in U(\beta,\eta,\zeta,\eta',\beta')$:

$$\mu((\varphi(x))_{a-N-\tau},\dots, \boldsymbol{(\varphi(x))_0},\dots,(\varphi(x))_{b+N+\tau})$$
$$= \int\chi((\varphi(x))_{a-N-\tau},\dots, \boldsymbol{(\varphi(x))_0},\dots,(\varphi(x))_{b+N+\tau})\, d\mu$$

$$= \int \chi(((\varphi(x))_{a-N-\tau}, \ldots, (\varphi(x))_0, \ldots, (\varphi(x))_{b+N+\tau}) \circ \varphi \, \exp \sum_{k \in \mathbf{Z}} -(f \circ \sigma^k \circ \varphi - f \circ \sigma^k) \, d\mu.$$

This leads to

(7) $\mu(\beta,\theta,\omega,\theta',\beta') \geq \mu(\beta,\eta,\zeta,\eta',\beta') \exp\{\sum_{k \in \mathbf{Z}} -(f \circ \sigma^k \circ \varphi(x) - f \circ \sigma^k(x)) - c\}$,

where

$$c = 4\|f\| \cdot \alpha^{\tau/2}.$$

In the last estimate we made use of the fact that

$$z_i = x_i = (\varphi(x))_i$$

for all $z \in U(\beta,\theta,\zeta,\theta',\beta')$ and $i \in [a-N-\tau, a-N] \cup [b+N, b+N+\tau]$. The same estimate holds on the shifted sets using φ'. For any $y \in U(\beta,\eta,\zeta^*,\eta',\beta')$

(8) $\mu(\beta,\theta,\omega^*,\theta',\beta') \geq \mu(\beta,\eta,\zeta^*,\eta',\beta') \exp\{\sum_{k \in \mathbf{Z}} -(f \circ \sigma^k \circ \varphi'(y) - f \circ \sigma^k(y)) - c\}$

holds true. Set $y = \sigma(x)$ and (7) and (8) combined to give

$$\mu(\beta,\theta,\omega^*,\theta',\beta') \geq \rho \cdot \mu(\beta,\theta,\omega,\theta',\beta') \exp\{\sum_{k \in \mathbf{Z}} [\,(f \circ \sigma^k \circ \varphi(x) - f \circ \sigma^k(x)) - (f \circ \sigma^k \circ \varphi' \circ \sigma(x) - f \circ \sigma^k \circ \sigma(x))] - 2c\}.$$

Set $M^* = \max(|a-N|, |b+N|)$, choose $M > M^*$ and estimate the sum in the exponential by

$$\leq \sum_{|k| \geq M} \{|f \circ \sigma^k \circ \varphi' \circ \sigma(x) - f \circ \sigma^{k+1}(x)| + |f \circ \sigma^k \circ \varphi(x) - f \circ \sigma^k(x)|\} + |\sum_{|k|<M} \{-(f \circ \sigma^k \circ \varphi' \circ \sigma(x) - f \circ \sigma^k \circ \varphi(x)\} + f \circ \sigma^M(x) - f \circ \sigma^{1-M}(x)|$$

and since we identify $\sigma \circ \varphi(x)$ with $\varphi' \circ \sigma(x)$, it follows

$$\leq 2\|f\| \cdot \alpha^{M+\tau-M^*}/(1-\alpha) + |-f \circ \sigma^{M-1} \circ \varphi' \circ \sigma(x) + f \circ \sigma^{1-M} \circ \varphi(x) + f \circ \sigma^M(x) - f \circ \sigma^{1-M}(x)|$$

$$\leq 2\|f\| \cdot (\alpha^{M+\tau-M^*}/(1-\alpha) + \alpha^{M-M^*-1}).$$

This tends to zero as M tends to infinity. By the choice of τ we obtain

$$\mu(\beta,\theta,\omega^*,\theta',\beta') \geq \sqrt{\rho} \cdot \mu(\beta,\theta,\omega,\theta',\beta')$$

for all $\omega \in \mathcal{T}[b-a]$ and suitable $\theta,\theta' \in \mathcal{T}[N-1]$. The strings β and β' are determined by the point $x \in U(\beta,\theta,\zeta,\theta',\beta')$, and therefore only a part of Σ_A

gets covered by varying ω over the whole $\mathcal{T}[b-a]$. Now we proceed to cover in the second generation, whereby we partly cover the complement of what was already covered in the first step. Let the strings β, η, ζ etc. now be denoted with an index 1 (β_1, η_1, ζ_1 etc.) and set $\zeta_2 = \beta_1,\eta_1,\zeta_1,\eta'_1,\beta'_1$. The cylinder $U(\zeta_2)$ now gets covered by smaller ones, and again there is at least one cylinder for which

$$\mu(\beta_2,\eta_2,\zeta_2^*,\eta'_2,\beta'_2) \geq \rho\cdot\mu(\beta_2,\eta_2,\zeta_2,\eta'_2,\beta'_2),$$

where $\beta_2,\beta'_2 \in \mathcal{T}[\tau]$ and $\eta_2,\eta'_2 \in \mathcal{T}[N-1]$. For the second generation ω_2, θ_2, θ'_2 we proceed as above. Call the union of all cylinders constructed in the i-th generation V_i for $i \in \mathbb{N}$, then we have $V_i \cap V_j = \emptyset$ if $i \neq j$. It is possible to cover in this manner arbitrarily large subsets of Σ_A. To make it obvious we will show that

$$\mu(\complement(\textstyle\bigcup_{1\leq i\leq n} V_i)) \to 0$$

for $n \to \infty$, where $\complement$ denotes the complement in Σ_A. In constructing the cylinders on which φ respectively φ' are defined we keep fixed the $(\tau+N)$-strings at each end of ω, and thus we select just one small cylinder from at most $(|A|\cdot\|A^{\tau+N}\|)^2$ small cylinders. Their μ-measures may be compared, for example by constructing conjugating homeomorphisms. Since μ is Gibbs the ratio of two of them is at most

$$\exp[2\cdot\|f\|(\tau + N + 2/(1-\alpha))].$$

This allows us to deduce a lower bound for the measure of V_n for $n \geq 1$, depending on all the previous generations these subsets. If we set

$$\Xi = (|A|\cdot\|A^{\tau+N}\|)^{-2}\exp[-2\cdot\|f\|(\tau + N + 2/(1-\alpha))],$$

which is a positive constant, we conclude that

$$\mu(V_n) \geq \Xi\cdot\mu(\complement(\textstyle\bigcup_{1\leq i<n} V_i)),$$

for all $n \in \mathbb{N}$. Since the V_i are pairwise disjoint, we have

$$\begin{aligned}\mu(\complement(\textstyle\bigcup_{1\leq i\leq n} V_i)) &= \mu(\complement(\textstyle\bigcup_{1\leq i<n} V_i)) - \mu(V_n)\\ &\leq (1 - \Xi)\mu(\complement(\textstyle\bigcup_{1\leq i<n} V_i)).\end{aligned}$$

Iteration yields

$$\mu(\complement(\textstyle\bigcup_{1\leq i\leq n} V_i)) \leq \mu(\complement V_1)(1 - \Xi)^{n-1}.$$

and this tends to zero as n tends to infinity. By construction it is

$$\mu(\sigma(V_i)) \geq \sqrt{\rho}\cdot\mu(V_i)$$

for all $i \in \mathbf{N}$, and hence

$$\mu(\sigma(\Sigma_A)) \geq \mu(\sigma(\bigcup_{1\leq i\leq n} V_i)) \geq \sqrt{\sqrt{\rho}} > 1$$

if n is large enough. This is impossible. In case $\mu(\sigma(B)) \leq \rho\cdot\mu(B)$ for $\rho \in (0,1)$ we replace σ by σ^{-1}, ρ by ρ^{-1} and proceed in the same way.

□

Proposition 11: *Let f be a function on Σ_A which decreases exponentially fast with $\alpha \in (0,1)$ and let σ be topologically mixing. If μ is a Gibbs state for f then it is also an equilibrium state.*

The **proof** is a reference to [1], Theorems 1.2 and 1.22. In Lemma 7 we have checked the conditions for a Gibbs state in Bowen's sense. Lemma 10 shows σ-invariance of μ and thus it is an equilibrium state for f on (Σ_A,σ). Since f is Hölder continuous it is the unique one.

□

Lemma 12. *Suppose a Smale space with a homeomorphism acting topologically mixing implies that a Gibbs state for a given function has to be an equilibrium state.*

Let ν be any Gibbs state. Then there exists a number $s \in \mathbf{N}$, such that ν is an equilibrium state for $F_s = \sum_{r\in[0,s)} F\circ T^r$ on (Ω,T^s).

The **proof** is by Smale's spectral decomposition (cf. [1], theorem 3.5). The non-wandering set of Ω is a union of finitely many disjoint compact sets Ω^v, called basic sets, which are invariant under T and on which T acts topologically transitively. Points which are conjugated lie always in the same basic set Ω^v, each of which is itself a union of $t[v]$ many disjoint, compact sets $\Omega^{v,u}$ on which $T^{t[v]}$ acts topologically mixing and where we have $T^{t[v]}(\Omega^{v,u}) = \Omega^{v,u}$, for all $u \in [1,t[v]]$. Each set $\Omega^{v,u}$ has positive distance from all the others, larger than δ, say. For any conjugating homeomorphism ψ defined on $U \subset \Omega$ there is

$$d(T^{kt[v]} \circ \psi(z), T^{kt[v]}(z)) < \delta,$$

for all $z \in U$ and where $|k|$ is big enough. That is, ψ restricted to $U \cap \Omega^{v,u}$ maps again into $\Omega^{v,u}$. So we restrict to maps ψ acting only on $\Omega^{v,u}$, replace T by $T^{t[v]}$ and F by $F_v = \sum_{0 \le r < t[v]} F \circ T^r$; then

$$\log g = \sum_{i \in \mathbf{Z}} (F \circ T^i \circ \psi - F \circ T^i)$$

$$= \sum_{i \in \mathbf{Z}} \sum_{0 \le r < t[v]} (F \circ T^{r+it[v]} \circ \psi - F \circ T^{r+it[v]})$$

$$= \sum_{i \in \mathbf{Z}} (F_v \circ T^{it[v]} \circ \psi - F_v \circ T^{it[v]}).$$

If ν is a Gibbs state for F on (Ω,T) then for each v,u the normalisation of ν restricted to $\Omega^{v,u}$ is a Gibbs state for F_v on $(\Omega^{v,u}, T^{t[v]})$ and vice versa. Take s to be the lowest common multiple of the numbers $t[v]$.

□

Proof of the theorem. It follows immediatly from the Lemmas 9 and 10 that ν is invariant under some power of T. As noted at the beginning, the conjugating maps as constructed in Lemma 7 give rise to a finite number of homeomorphisms on (Ω,T). Finally, it is well-known that an equilibrium state on the shift space Σ_A corresponds automatically to one on Ω (cf. [4] Theorem 7.9).

□

We cannot expect ν to be an equilibrium state on (Ω,T), since that would require it to be T-invariant. A look at the spectral decomposition as described in Lemma 12 shows that this is in general not true. The measure ν restricted and normalised to $\Omega^{v,u}$ (if $\nu(\Omega^{v,u}) > 0$) is clearly Gibbs and is invariant under $T^{t[v]}$ but not under T unless $t[v] = 1$. But T need not be mixing to have an invariant measure. We have

Corollary 13: (i) *A* T *-invariant Gibbs measure for a Hölder continuous real-valued function over a Smale space is an eqilibrium state;*
(ii) *if* T *acts topologically mixing, then, by Lemma 10, a Gibbs measure for a Hölder continuous real-valued function over a Smale space is an equilibrium state.*

I am indebted to Peter Walters for encouragement, advice and especially for pointing out Lemma 5.

References:

[1] R. Bowen; *Equilibrium states and the ergodic theory of anosov diffeomorphisms.* SLN # 470, 1975.

[2] D. Capocaccia; A definition of Gibbs states for a compact set with $\mathbf{Z}^{\nu}$-action. Comm. Math. Phys. 48, (1976), pp. 85 - 88.

[3] A. Manning; Axiom A diffeomorphisms have rational Zetafunctions. LMS Bulletin, 3 (1971), pp. 215 - 220.

[4] D. Ruelle; *Thermodynamic Formalism.* Addison-Wesley, 1978.

An equivalence relation on shifts of finite type.

Abstract: On subshifts which are derived from Markov partitions exists an equivalence relation which identifies points that lie on the boundary set of the partition. In this paper we restrict to symbolic dynamics. We express the quotient space in terms of a non transitive subshift of finite type, give a necessary and sufficient condition for the existence of a local product structure and evaluate the Zeta function of the quotient space. Finally we give an example where the quotient space is again a subshift of finite type.

1. Introduction.

In this paper we are concerned with a special kind of equivalence relations which occur in symbolic dynamics. The questions treated here arise from Markov partitions of Axiom A diffeomorphisms (cf. Smale [7] p.777 and Bowen [1]). For a small enough partition one gets a shiftspace and a projection onto the original manifold where the diffeomorphism acts conjugate to the shift. It is known that a subshift of finite type can be isomorphic only to an Axiom A diffeomorphism over a non-wandering set of zero dimension. It is therefore clear that the boundary set, i.e. the set of points whose pre images in the shift consists of more than one point, containes essential information about the structure of the non wandering set despite the fact that it has measure zero for any smooth measure.

We begin this paper by demonstrating that it is enough to consider strings of some certain length whenever we want to decide whether a relation induces an equivalence relation on Σ_A. In the following five sections we restrict to equivalence relations that have finite equivalence classes. In that case the quotient space can be described by means of a non-transitive subshift, which has a partial ordering. Maximal elements with respect to this ordering correspond to points in the quotient space. This formulation will be used in section 3 to express the topology on the quotient space in terms of cylinder sets of this new shift space. In section 4 we reduce each shiftspace with such an equivalence relation to the case where transitive points have no other equivalent point except themself. In section 5 we shall give a necessary and sufficient condition for the existence of a local product structure on the quotient space. In section 6 we evaluate the Zeta function under the assumption made that the equivalence classes are finite. It turns out, that in this more general context the Zeta function can be evaluated by Mannings product formula (see [5]). Finally, in the last section we give an example of a shift space, for which the quotient space is again a subshift of finite type independent of the equivalence relation.

Let (Ω,T) be a Smale space (a compact metric space with an expanding transformation and a local product structure; see [6] p. 125 ff, and section 3) with metric $d(\cdot,\cdot)$ and homeomophism T with expansive constant ε. Let $\mathfrak{R} = \{R_j;\ j \in \mathfrak{J}\}$, for some index set $\mathfrak{J}$, be a Markov partition. Two rectangles R_i and R_j are said to be related if they have non-empty intersection, write $R_i \sim R_j$. We will say relation whenever we mean it to be symmetric and reflexive but not necessarily transitive. Let A be the 0-1-transition matrix associated to the Markov partition and set for the shift space

$$\Sigma_A = \{(x_i)_{i\in\mathbf{Z}}:\ A[x_i,x_{i+1}] = 1 \text{ for all } i \in \mathbf{Z}\}.$$

Define on Σ_A the relation $\approx$ as: $x \approx y$ if and only if $x_i \sim y_i$ for all $i \in \mathbf{Z}$. In the case of a Markov partition with diameter of each rectangle less than $\varepsilon/2$, $\approx$ is transitive and therefore an equivalence relation since for any two points $x, y \in \Sigma_A$ with $x \approx y$ we have $\pi(x) = \pi(y)$. This follows from $d(T^k \circ \pi(x), T^k \circ \pi(y)) \le \varepsilon\ \ \forall\ k \in \mathbf{Z}$ using the expansiveness of T. The next lemma applies to an arbitrary Σ_A and relation on A. Let $\alpha = |\mathfrak{R}|^3$.

Lemma 1: A relation $\sim$ on $\mathfrak{R}$ induces an equivalence relation on Σ_A in the above manner if and only if for any three Σ_A-words $x_{-\alpha}\dots x_\alpha$, $y_{-\alpha}\dots y_\alpha$, $z_{-\alpha}\dots z_\alpha$ satisfying $x_i \sim y_i$, $y_i \sim z_i$ $\forall\ |i| \le \alpha$, one has $x_0 \sim z_0$.

Proof. First let us assume that every tripel of strings that are long enough implies transitivity of $\sim$ on the sections cut off on 'both sides' by α. Then it is clear that $\approx$ is an equivalence relation since the criterion applies to any finite section.

Secondly, suppose $\sim$ induces an equivalence relation $\approx$ on Σ_A, and there were three words of length $2\alpha+1$, $x_{-\alpha}...x_\alpha$, $y_{-\alpha}...y_\alpha$, $z_{-\alpha}...z_\alpha$ satisfying $x_i \sim y_i \sim z_i$ for $|i| \leq \alpha$ but for which we have as well $x_0 \not\sim z_0$. We shall contradict the transitivity of $\approx$. The strings are chosen long enough so that a tripel (y_k,x_k,z_k) appears twice on the positive side, i.e. we can find indices $0 \leq k < l \leq \alpha$ so that $(y_k,x_k,z_k) = (y_l,x_l,z_l)$. Iterating this loop yields three positively equivalent points. The same argument applied to negative indices and we obtain three points $x,y,z \in \Sigma_A$ for which $x \approx y \approx z$ holds true, but not $x \approx z$ since we have by construction $x_0 \not\sim z_0$. Hence the lemma. □

From this lemma we immediatly deduce the following result.

Corollary 2: Let $\approx$ be an equivalence relation and suppose there are three strings $x_k...x_l$, $y_k...y_l$ and $z_k...z_l$ for some k,l, $k+2\alpha < l$ for which $x_i \sim y_i \sim z_i$, $i = k,...,l$ holds true. Then also $x_i \sim z_i$ for $i = k+\alpha,...,l-\alpha$. □

2. A non transitive subshift.

From now on we assume $\approx$ to be an equivalence relation on Σ_A. Two strings $x_k...x_l$, $y_k...y_l$ $(k < l)$ in Σ_A are said to form a diamond if $x_k = y_k$, $x_l = y_l$ and a collapsing diamond if additionally $x_i \sim y_i$ for $i = k,...,l$. Furthermore we will assume the equivalence classes in Σ_A to be finite, which is the same as to demand that there are no collapsing diamonds. That is the situation we are in, if Σ_A is a Markov shift and the partition is fine enough (i.e. if the diameter of the rectangles covering Ω are less than half of an expansive constant). We refer the reader especially to [1] chapter 2. To prove that the projection π: $\Sigma_A \to \Omega$ is bounded to one shows in effect, that Σ_A cannot have collapsing diamonds and π can therefore be at most [illegible]^2 to one. From this it seems reasonable to turn the argument into an assumption, and that is what we have done here. To verify that $\approx$ does not collapse diamonds it is enough to check all possible pairs of strings with length of at most $|\mathfrak{A}|^2+1$. In particular, as already remarked, the equivalence classes contain at most $|\mathfrak{A}|^2$ elements. Furthermore we assume Σ_A is topologically mixing, i.e. A^n is positive for n large enough.

Let $\mathfrak{A}_n$ denote the collection of the (unordered) n-element subsets $\{a_1,\dots,a_n\}$ of $\mathfrak{A}$, which contain no symbol twice and satisfy $a_i \sim a_j$ for any two i,j in $[1,n]$. Introduce an ordering on $\mathfrak{A}_n$ in the following way: $\{a_1,\dots,a_n\} \leq \{b_1,\dots,b_n\}$ if there are admissible strings of some length $m+1$: $x^i_0 \dots x^i_m$ with $x^i_0 = a_i$, $x^i_m = b_i$ for $i = 1,\dots,n$, and $x^i_k \sim x^j_k$ for any two i,j in $[1,n]$ and $k = 0,\dots,m$. This generates a decomposition of $\mathfrak{A}_n$ into subsets $\mathfrak{A}_n^k$ for $k = 1,\dots,k_n$, in the way that for any two elements ξ_0,ξ_1 in the same $\mathfrak{A}_n^k$ we have it both ways $\xi_0 \leq \xi_1$ and $\xi_1 \leq \xi_0$ and in particular $\xi_0 \leq \xi_0$. For elements ξ and ζ in different $\mathfrak{A}_n^k$ we have either $\xi < \zeta$ or $\zeta < \xi$ if these two subsets may be joined up at all (where we said $<$ if $\leq$ holds but not $\geq$). This ordering process can be extended to the entire collection $\{\mathfrak{A}_n^k\colon k,n\}$. Two subsets $\mathfrak{A}_n^k$, $\mathfrak{A}_m^l$ are either related (e.g. $\xi \leq \zeta$ for some $(\xi,\zeta) \in \mathfrak{A}_n^k \times \mathfrak{A}_m^l$), i.e. $\mathfrak{A}_n^k < \mathfrak{A}_m^l$ or $\mathfrak{A}_m^l < \mathfrak{A}_n^k$ for $(n,k) \neq (m,l)$, or no element in $\mathfrak{A}_n^k$ may be connected via some strings in the described manner with $\mathfrak{A}_m^l$ (or the other way round). In the last case $\mathfrak{A}_n^k$ and $\mathfrak{A}_m^l$ are said to be separated. Furthermore, it is impossible that there are $\xi_0,\xi_1 \in \mathfrak{A}_n^k$ and $\zeta \in \mathfrak{A}_m^l$ for some $(n,k) \neq (m,l)$, with $\xi_0 \leq \zeta \leq \xi_1$. The case $n = m$ is clear, so let $n \neq m$. Since $\xi_1 \leq \xi_0$, the chain extends to $\xi_0 \leq \zeta \leq \xi_1 \leq \xi_0$, which means that there are strings beginning in ξ_0 and returning to it. A string that begins in ξ_0 on a particular element does not necessarily end up on again the same element. Hence, a set of strings that materializes $\xi_0 \leq \zeta \leq \xi_1 \leq \xi_0$, induces a permutation π on the elements of ξ_0. Some power of π is the identity permutation and would imply collapsing diamonds since ξ_0,ξ_1 on the one hand and ζ on the other have different cardinality (because $n \neq m$). For $n \neq m$ we have thus always either $\mathfrak{A}_n^k < \mathfrak{A}_m^l$ or $\mathfrak{A}_m^l < \mathfrak{A}_n^k$ if they are not separated.

Consider $\{\mathfrak{A}_n^k\colon k,n\}$ as a new alphabet and denote it by $\mathfrak{C}$. Delete all $\mathfrak{A}_n^k$ which may not be extended infinitly forward and backward, i.e. all these elements for which all possible transitions forward or backward inevitably lead to a dead end (For the same effect one can also pass to higher block systems). Notice that in a subalphabet $\mathfrak{A}_n^k$ that containes at least two symbols, every sumbol ξ may at some future time be followed up by it itself, i.e. $\xi \le \xi$. Define a transition matrix C for the alphabet $\mathfrak{C}$ by setting $C[\xi,\zeta] = 1$ if there exist Σ_A-words $\{a_i a_j;\ i,j\}$ of length 2 so that $\{a_i\} = \xi$ and $\{a_j\} = \zeta$, where $\xi,\zeta \in \{\mathfrak{A}_n^k:\ k,n\}$; and $C[\xi,\zeta] = 0$ otherwise. This defines in the usual way a (non transitive subshift) Σ_C of type two, where, with a suitable ordering of $\mathfrak{C}$, the transition matrix C is of diagonal form

$$\begin{vmatrix} C(1) & 0 \ldots\ldots & 0 \\ 0 & C(2) & 0 \ldots 0 \\ : & \ldots\ldots\ldots & : \\ 0 \ldots\ldots\ldots & 0 & C(s) \end{vmatrix},$$

for some s, where the submatices are square of upper triangular form

$$\begin{vmatrix} A(t,1) & * \ldots\ldots & * \\ 0 & A(t,2) & * \ldots * \\ : & \ldots\ldots\ldots & : \\ 0 \ldots\ldots\ldots & 0 & A(t,r) \end{vmatrix},$$

for some $r = r_t$. We call the subshifts $\Sigma_{C(i)}$ the components of Σ_C. Among the C(i) is one contains the transitionmatrix A somewhere in the diagonal. We call that particular sub-C and the associated shift space as the principal component of Σ_C. Call $\mathfrak{C}_i$ the subalphabets which determin the subshifts $\Sigma_{A(k,l)}$ and arrange the indices (decreasing, so that $\mathfrak{C}_i > \mathfrak{C}_j$ if $i > j$, or $\mathfrak{C}_i$ and $\mathfrak{C}_j$ are not separated) so that $\mathfrak{C}_0 = \mathfrak{A}$ Accordingly the matrices A(k,l) are reindexed A(i) (with A(0) = A). Observe

(i) that every subset of the form $\{a_1,\ldots,a_n\}$ with $a_i \sim a_j$ for any i,j in [1,n] appears in exactly one of the $\mathfrak{C}_i$;

(ii) the alphabet $\mathfrak{C}$ is closed under intersections of its elements regarded as sets;

(iii) the subshifts $\Sigma_{A(i)}$ are topologically transitive if not empty (Such a subshift is over a one-element subalphabet consisting of some ξ for which $\xi \le \xi$ does not hold true.);

(iv) there is an integer function $\vartheta(i)$ which denotes the cardinality of the elements in the alphabet $\mathfrak{C}_i$, the number $\aleph(i) = \gamma(i)-1$ is the <u>dimension</u> of $\mathfrak{C}_i$.

We conclude this section by listing some properties of Σ_C. To begin with, we define a matrix $C^{\sim}$ of the same size as C and which has zeros at exactly the same places as C. Let $\xi = \{a_1,...,a_n\}$, $\zeta = \{b_1,...,b_m\}$ elements of $\mathfrak{C}$, the entries $C^{\sim}[\xi,\zeta]$ then are defined as the number of different sets of 2-strings of the form a_ib_j, $a_i, b_j \in \mathfrak{A}$, that materialize the transition $\xi \to \zeta$. Here are some examples to illustrate $C^{\sim}$:

$C^{\sim}[\{a,b,c\},\{a',c'\}] = 1$ if $(\{a,b\}\mapsto a', c\to c')$;

$C^{\sim}[\{a,b\},\{a',b'\}] = 2$ if $(\{a,b\}\mapsto a', \{a,b\}\mapsto b')$;

Lemma 3: Let $\xi_0...\xi_k$ be a word in Σ_C, with all ξ_i in the same subalphabet C_l. Then there is exactly one set of related Σ_A-strings $x^s{}_0...x^s{}_k$, $s = 1,..., \gamma(l)$, with $\xi_i = \{x^1{}_i,...,x^{\gamma(l)}{}_i\}$ and $x^s{}_i \sim x^{s'}{}_i$ for any two $s,s' \in [1,\gamma(l)]$ and $i = 0,...,k$.

Proof. We have to show that $C^{\sim}[\xi,\zeta]$ is at most one for any two ξ and ζ belonging to the same subalphabet. Suppose the statement were false then $C^{\sim}[\xi,\xi] \geq 2$ fore some ξ,ζ with $\xi \leq \zeta \leq \xi$. Set $\xi = \{a_1,...,a_n\}$, $\zeta = \{b_1,...,b_n\}$, fix an ordering of the elements and select under possibly several transitions one linking ζ to ξ. We end up having the situation $\xi \to \zeta \to\to \xi$, with two different transitions in the first position (related to each other because they consist of only two elements) together with a unique transition in the second place (fixed once chosen). Hence $C^{\sim}[\xi,\xi] \geq 2$ induces on ξ two different permutations π_1 and π_2. There are integers $k,l \geq 1$, so that $\pi_1{}^k = \pi_2{}^l = \mathbb{1}$ and by iterating the loop $\xi \to \zeta \to\to \xi$ kl-times we get collapsing diamonds, since $\pi_1{}^{kl} = \pi_2{}^{lk} = \mathbb{1}$ on ξ. □

Given x in Σ_A, then we set $\langle x\rangle = \{z\colon z \in \Sigma_A$ such that $z \approx x\}$ for the equivalence class of x. Analoguously, we write $\langle x\rangle_i = \{z_i\colon z \in \langle x\rangle\}$ for the collection of the i-th coordinates as a subset of $\mathfrak{A}$. Clearly $\langle x\rangle$ is a point in Σ_C.

Lemma 4: Let x be a periodic point in Σ_A. Then $\langle x\rangle$ lies in a subshift $\Sigma_{A(l)}$ for some l.

Proof. Assume $\langle x\rangle$ does not lie entirely in some $\Sigma_{A(l)}$. However, for some $N \in \mathbb{Z}$ we have $\langle x\rangle_i \in \mathbb{C}_k \ \forall\ i \le N$ and some k. The $\langle x\rangle_i$ themself are periodic for $i \le N$ (not necessarily with the same period as x) and therefore may be extended to a periodic point $x^{\sim} \in \Sigma_C$ with $x^{\sim}_i \in \mathbb{C}_k$ for all $i \in \mathbb{Z}$ and with $x^{\sim}_i = \langle x\rangle_i$ for $i < N$. All Σ_A-strings running through $x^{\sim}$ are necessarily equivalent to x (by periodicity) from which follows by corollary 2 that $x^{\sim}_i \subset \langle x\rangle_i$ for all $i \in \mathbb{Z}$. By the same argument we have $\langle x\rangle_i \in \mathbb{C}_l \ \forall\ i \ge M$ some integer M and some $\mathbb{C}_l$. In the same way one defines a point x^* in Σ_C with $x^*_i \in \mathbb{C}_l$ for all integers i so that $x^*_i = \langle x\rangle_i$ for $i \ge M$. In a similar way we conclude $x^*_i \subset \langle x\rangle_i$ for all $i \in \mathbb{Z}$ and hence $x^{\sim} = x^* = \langle x\rangle$. □

On Σ_C there is a partial ordering by inclusion. For $x,y \in \Sigma_C$ we say $x \subset y$ if $x_i \subset y_i$ as sets for all integers i. For x out of Σ_A, $\langle x\rangle$ are maximal elements in Σ_C and vice versa, maximal elements in Σ_C correspond to points in the quotient $\Sigma_A/\approx$.

Lemma 5: Given x a periodic point in Σ_A, then $\langle x\rangle$ has the same period in Σ_C.

Proof. Let n be the (least) period of x an suppose $\langle x\rangle$ has a period which is a multiple of n, rn say. Then $\langle x\rangle_0 \neq \langle x\rangle_{sn}$ for some $0 < s < r$, and since x is a string running through $\langle x\rangle$ we have $x \in \bigcap_{0\le s<r} \sigma^{sn}\langle x\rangle$. Thus we define a new sequence $x^* = \{x^*_i\!: \ i \in \mathbb{Z}\}$ by setting $x^*_i = \bigcup_{0\le s<r} \langle x\rangle_{i+sn}$. Then $\langle x\rangle \subset x^*$ and $\langle x\rangle \neq x^*$; furthermore x is a sequence contained in x^* and all other sequences are equivalent to x and, by virtue of corollary 2, are equivalent to each other. This shows that x^* is a point in Σ_C which contains $\langle x\rangle$, therefore $x^* = \langle x\rangle$ since $\langle x\rangle$ is maximal. □

The last lemma is not true when one fixes an ordering on the symbols of $\mathbb{C}$

3. The topology on $\Sigma_A/\approx$.

As in corollary 2 let $\alpha =$ [illegible], and denote $\Sigma_A/\approx$ by Ω. Define $U_n = \{(x,y) \in \Sigma_A \times \Sigma_A\colon\ x_i \sim y_i \text{ for all } |i| \le 2\alpha n\}$. Clearly, each U_n contains the diagonal of $\Omega\times\Omega$ and is symmetric. Furthermore for each n we have $U_n \circ U_n \circ U_n \subset U_{n-1}$, where $U_n \circ U_n = \{(x,z)\colon\ \exists\ y \in \Sigma_A$ such that $(x,y), (y,z) \in U_n\}$. To see this, choose points $w, x, y, z \in \Sigma_A$ with $w_i \sim x_i \sim y_i \sim z_i$ for all $|i| \le 2\alpha n$. Corollary 2 applied gives firstly $w_i \sim y_i$ for $|i| \le \alpha(2n-1)$ and secondly $w_i \sim z_i$ for all $|i| \le 2\alpha(n-1)$. Hence $U_n \circ U_n \circ U_n \subset U_{n-1}$ holds true for all $n \in \mathbb{N}$, and we may therefore apply Frinke's metrization lemma (see [4] p. 185) which says that there exists a pseudo-metric d on Ω, with the property $U_n \subset \{(x,y)\colon\ d(x) < 2^{-n}\} \subset U_{n-1}$ for all integers n. In fact, d is a metric since $x,y \in \Sigma_A$ represent the same point in Ω if and only if $x_i \sim y_i\ \ \forall\ i \in \mathbb{Z}$, which is the case if (x,y) lies in U_n for all n. See also D. Fried [3].

One would like to consider as distance function d^*, for $d^*(x,y) = \lambda^p$ where $p = p(x,y) = \max\{q\colon\ x_i \sim y_i\ \ \forall\ |i| \le q\}$ for some $\lambda \in (0,1)$. Indeed, for $\lambda = 2^{-2\alpha}$, d^* is equivalent to d: $C^{-1}d(.,.) \le d^*(.,.) \le d(.,.)$, where $C = 2^{2\alpha}$. Set $\beta = |2$[illegible]$|$ and for later use we deduce the following lemma

Lemma 6: Let x, y, w be three points in Σ_A satisfying (i) $x_i \sim y_i$ for $i \le k$ and (ii) $w_i \sim x_i$ for $i \le k-\beta$. Then $w_i \sim y_i$ for $i \le k-\beta$.

Proof. Let us assume $k-\beta$ is positive and a multiple of 2α, $k-\beta = 2n\alpha$ for some $n > 1$, say. We have therefore that (x,w) lies in U_{n-1}, and since $U_{n-1} \subset \{(x,y)\colon\ d(x,y) < 2^{1-n}\} \subset U_{n-2}$ it follows that $d(x,w) < 2^{1-n}$. By the same argument we obtain $d(x,y) < 2^{-n}$. We shall construct two equivalent points x', y'. Set $(x'_i, y'_i) = (x_i, y_i)$ for $i \le k-\beta$. Since $\beta = |2$[illegible]$|$ there is an index $i \in (k-\beta, k]$ so that $\{x_i, y_i\}$ lies in some $\mathbb{C}_j$ for which the associated shift space $\Sigma_{A(j)}$ is transitive. Now choose any half infinite word in $\Sigma_{A(j)}$, $\{x'_k, y'_k\}\{x'_{k+1}, y'_{k+1}\}\{x'_{k+2}, y'_{k+2}\}\ldots$, satisfying $\{x'_k, y'_k\} = \{x_k, y_k\}$. Since d is a metric on $\Sigma_A/\approx$ we have $d(y',w) \le d(y',x') + d(x',w) \le d(z',w) < 2^{1-n}$ which implies $y'_i \sim w_i \ \forall \ i \le k-\beta$, and because y, w did not get changed on positions less or equal to $k-\beta$ it follows $y_i \sim w_i \ \forall \ i \le k-\beta$. If k is an arbitrary integer we use a suitable iterate of σ to bring x,y,w into a position so that the new $k-\beta$ is positive and a multiple of 2α. Hence the lemma. □

It is easily seen that the lemma holds true too in the right asymptotic case: If $x,y,w \in \Sigma_A$ which satisfy (i) $x_i \sim y_i \ \forall \ i \ge k$ and (ii) $w_i \sim x_i \ \forall \ i \ge k+\beta$, for some $k \in \mathbf{Z}$ then $w_i \sim y_i \ \forall \ i \ge k+\beta$. A consequence of this lemma is the following statement:

Proposition 7: Suppose $x_k...x_l \sim y_k...y_l \sim z_k...z_l$ are three strings, which are related in the way indicated, where $l-k$ is at least $\beta+2\alpha$, and $\{x_k,y_k\}$, $\{x_l,y_l\}$ are symbols in $\mathbf{C}$. Then $x_k...x_l$ is related to $z_k...z_l$.

Proof. By corollary 2 we know already that $x_i \sim z_i$ for $i = k+\alpha,...,l-\alpha$. By assumption the strings $x_{k+\alpha}...x_{l-\alpha}$, $y_{k+\alpha}...y_{l-\alpha}$, $z_{k+\alpha}...z_{l-\alpha}$ have length at least β. We may therefore construct three points $x',y',z' \in \Sigma_A$ with the properties

(i) $(x'_i,y'_i,z'_i) = (x_i,y_i,z_i)$ for $i = l-\alpha+1,...,l$,

(ii) $x_i \sim y_i$, $x_i \sim z_i$, $y_i \sim z_i$ for $i \le l-\alpha$,

(iii) the symbols x'_i, y'_i for $i > l$ are choosen to make x' and y' equivalent points which is possible by assumption; and z'_i for $i > l$ may be anything in Σ_A.

We have now the situation $x'_i \sim y'_i \sim z'_i \ \forall \ i \le l$ and $x' \approx y'$. Using Lemma 6 we conclude that $x_i \sim z_i$ for $i = l-\alpha+1,...,l$. In the same way one constructs three right assymptotic points and proves $x_i \sim z_i$ for $i = k,...,k+\alpha-1$. □

As remarked in the previous section, there is an ordering on Σ_C. The same applies to finite strings: we say $\mathfrak{s}_k...\mathfrak{s}_l \subset \mathfrak{m}_k...\mathfrak{m}_l$ if $\mathfrak{s}_i \subset \mathfrak{m}_i$ for $i = k,...,l$. Intersections of strings are defined in the obvious way. A basis for the topology induced by d on Ω is the set of all cylinders $U(\mathfrak{s}_k...\mathfrak{s}_l) = \{z \in \Sigma_A: \ z_i \in \mathfrak{s}_i, \ k \le i \le l\}$, where $\mathfrak{s}_k...\mathfrak{s}_l$ are finite strings in Σ_C.

For $x \in \Sigma_A$ define:

$$W^s(x,k) = \{z \in \Sigma_A\colon\ z_i \sim x_i \quad \forall\ i \geq -k\},$$

$$W^u(x,k) = \{z \in \Sigma_A\colon\ z_i \sim x_i \quad \forall\ i \leq k\}.$$

The union over k turns out to be the stable, $W^s(x)$, respectively unstable, $W^u(x)$, direction through x. The shift σ on Σ_A induces a homeomorphism on Ω which we again denote by σ. Pick a $z \in W^s(x,k)$ for some k so that $d^*(x,z) \leq d(x,y) \leq 1/2$. The homeomorphism σ acts on $W^s(x,k)$ therefore contracting distances d^* by λ and σ^{-1} contracts distances on $W^u(x,k)$ by a factor λ. Hence the stable and unstable directions through the points of Ω are described exactly by $W^s(x)$ and $W^u(x)$. Clearly, periodic points are dense in Ω.

4. Reducing Σ_A.

A point x in Σ_A $(\Sigma_{A(l)})$ is (doubly) transitive if for every $y \in \Sigma_A$ $(\Sigma_{A(l)})$ and $n \in \mathbb{N}$ there are positive integers m, m^*, so that $y_i = (\sigma^m x)_i = (\sigma^{m^*} x)_i$ for all $|i| \leq n$. In other words, every Σ_A-word ($\Sigma_{A(l)}$-word) appears infinitly often in the past and future dimensions of x. In this section we discuss the possibility that Σ_A may have transitive points with non-trivial equivalence classes. We shall show that Σ_A than can be replaced by another subshift of finite type in which transitive points have trivial equivalence classes and whose quotient is isomorphic to $\Sigma_A/\approx$.

For a $\Sigma_{A(l)}$-word $\chi_s\chi_{s+1}\cdots\chi_t$ we set $U(\chi_s\chi_{s+1}\cdots\chi_t) = \{\xi \in \Sigma_A\colon \xi_s\xi_{s+1}\cdots\xi_t = \chi_s\chi_{s+1}\cdots\chi_t\}$ for the cylinder of all points in $\Sigma_{A(l)}$ which have the word $\chi_s\chi_{s+1}\cdots\chi_t$ on the places between s and t. For a positive integer k we denote by τ^k the concatenation $\tau\tau\ldots\tau$, k-times.

We now pass to a higher block system. Without changing the notation we replace $\mathfrak{A}$ by the set of all $(\beta+2\alpha)$-words. The new transition matrix $A^\wedge$ is defined by setting $A^\wedge[x_1x_2\ldots x_{\beta+2\alpha}, y_1y_2\cdots y_{\beta+2\alpha}] = 1$ if $x_1x_2\ldots x_{\beta+2\alpha}, y_1y_2\cdots y_{\beta+2\alpha}$ are Σ_A-words satisfying $x_i = y_{i-1}$ for $i = 2,\ldots,\beta+2\alpha$; and 0 otherwise. The non-transitive subshift constructed in section 2 is now thought as being derived from this $(\beta+2\alpha)$-system (without introducing new notation). Naturally, there exists an induced relation on the new alphabet $\mathbb{C}^\wedge$: For $\xi,\zeta \in \mathbb{C}^\wedge$, we say $\xi \sim \zeta \Leftrightarrow a \sim b$ for some $(a,b) \in \xi\times\zeta$. Set $(a,b) = (a_1 \dots a_{\beta+2\alpha}, b_1 \dots b_{\beta+2\alpha})$; since $\{a_1,b_1\}$, $\{a_{\beta+2\alpha}, b_{\beta+2\alpha}\}$ are pairs of related symbols that can be prolonged infinitely in backward respectively forward direction (i.e. are elements of the ancient $\mathbb{C}$) we conclude, using proposition 7, that the induced relation on the higher block alphabet $\mathbb{C}^\wedge$ reads in fact: $\xi \sim \zeta \Leftrightarrow a \sim b$ for **all** $(a,b) \in \xi\times\zeta$. It is clear that the relation induces an equivalence relation on the subshift $\Sigma_{A^\wedge}$. From now on we shall call $(\mathbb{S}^\wedge, \Sigma_{A^\wedge}, \mathbb{C}^\wedge, \Sigma_{C^\wedge})$ by $(\mathbb{S}, \Sigma_A, \mathbb{C}, \Sigma_C)$, and the same applies to $A(l)$ etc.. Equivalence classes in $\Sigma_{A(l)}$ are denoted by $\langle . \rangle$, in the same way as in $\Sigma_{A(l)}$.

Lemma 8: Let χ be a transitive point in some $\Sigma_{A(l)}$. Then two sequences $\xi,\zeta \in \langle\chi\rangle$ are either identical or disagree on all places.

Proof: Let χ be a transitive point in Σ_A. We have to show that different points $\xi,\zeta \in \langle\chi\rangle$ differ on all places, i.e. $\xi_i \neq \zeta_i \ \forall \ i \in \mathbb{Z}$. Suppose there exists a $l \in \mathbb{Z}$ so that $\xi_l = \zeta_l$ and let $k \in \mathbb{Z}$ be an integer such that $\xi_k \neq \zeta_k$. Since χ is transitive, there are numbers $s,t \in \mathbb{Z}$, $s \le \min(k,i) < \max(k,i) \le t$, so that for any other transitive point $\chi' \in U(\chi_s\chi_{s+1}\cdots\chi_t)$ there exist $\xi',\zeta' \in \langle\chi'\rangle$ satisfying $(\xi'_j,\zeta'_j) = (\xi_j,\zeta_j) \ \forall \ \min(k,i) \le j \le \max(k,i)$. (Note, that the map $\chi \to \langle\chi\rangle_0 \in \mathbb{C}$ is continuous but not uniformly continuous.) In particular, since χ is transitive, the word $\chi_s\chi_{s+1}\cdots\chi_t$ appears infinitly often, say at intervals of length $m[1]$, $m[2]$,... (all bigger than $t-s$). Unfortunatly $(\xi_k \cdots \xi_l \cdots \xi_{k+m[1]}, \xi_k \cdots \zeta_l \cdots \xi_{k+m[1]})$ need not to have collapsing diamonds in Σ_A; this is because strings beginning in ξ_k on the same element do not necessarily end up again on the same element in $\xi_{k+m[1]}$. However, since the $\xi_{k+m[i]}$ consist of finitely many sumbols of $\mathbb{S}$, we can find two numbers $p \le q$, so that $(\xi_{k+m[p]} \cdots \xi_{k+m[q]}, \xi_{k+m[p]} \cdots \xi_{k+m[q]})$ collapses Σ_A-diamonds. □

Let $\mathbb{C}_k$ and $\mathbb{C}_l$ be two subalphabets satisfying $\partial(k) = \partial(l)+1$ and set $\partial(l) = m$. Denote by $\mathbb{C}_l^*$ the power-set of $\mathbb{C}_l$ and let us define a map $\upsilon: \mathbb{C}_k \to \mathbb{C}_l^*$ as follows: For $\{\chi^0,\chi^1,\dots,\chi^{m+1}\} = \chi \in \mathbb{C}_k$ we set

$$\upsilon(\chi) = \{\{\chi^1,...,\chi^{m+1}\}, \{\chi^0,\chi^2,...,\chi^{m+1}\},..., \{\chi^0,\chi^1,...,\chi^m\}\} \cap \mathbb{C}_l.$$

Lemma 9: Suppose $\chi,\chi' \in \mathbb{C}_k$ and $\chi \to \chi'$. Then for every $\zeta \in \upsilon(\chi)$ there exists exactly one $\zeta' \in \upsilon(\chi')$ so that $\zeta \to \zeta'$.

Proof. (i) Choose $\chi \to \chi'$ for some $\chi,\chi' \in \mathbb{C}_k$. It is then clear that every $\zeta \in \upsilon(\chi)$ has at least one successor in $\upsilon(\chi')$, namely the unique subset of χ which follows ζ when we join χ and χ' up by Σ_A-strings of length 2.

(i) Suppose there were two different ξ_0,ξ^*_0 in $\upsilon(\chi')$ and $\zeta \in \upsilon(\chi)$ for some $\chi,\chi' \in \mathbb{C}_k$, $\chi \to \chi'$, so that $\zeta \to \xi_0$, $\zeta \to \xi^*_0$. We shall construct collapsing diamonds. Since $\Sigma_{A(k)}$ is transitive we have $\chi \le \chi' \le \chi$ and therefore a string $\chi \to \chi' \to \chi_1 \to ... \to \chi_s \to \chi$ of some length s, where the χ_t are symbols in $\mathbb{C}_k$. Then we have $\zeta \to \xi_0 \to \xi_1 \to ... \to \xi_s \to \zeta$ for a sequence $\{\xi_t \in \mathbb{C}_l\colon \xi_t \subset \chi_t,\ t = 1,...,s\}$. On the other hand, there exists a similar sequence running through ξ^*_0 and returning to it, namely $\zeta \to \xi^*_0 \to \xi^*_1 \to ... \to \xi^*_s \to \zeta^*_{s+1} \to \xi^*_0$ where the ξ^*_t are symbols in $\mathbb{C}_l$ and subsets of the χ_t. Both, the ξ_t and the ξ^*_t, are subsets of the same χ_t, but unfortunately $(\zeta\xi_0... \xi_s\zeta\xi^*_0, \zeta\xi^*_0... \xi^*_s\xi^*_0)$ need not to provide collapsing diamonds in Σ_A, since strings that begin in ζ on the same element do not necessarily share again the same element in ζ^*_0. This difficulty is overcome by iterating the words $\zeta\xi_0... \xi_s$, $\xi^*_1... \xi^*_s\xi^*_0$ some $p \ge 1$ times; enough to make $\pi^p, \pi^{*p} = \mathbb{1}$, where π, π^* are the permutation on the element of ζ respectively ξ^*_0 which are induced by the Σ_A-strings running through $\zeta\xi_0... \xi_s\zeta$, $\xi^*_0\xi^*_1... \xi^*_s\xi^*_0$. This finishes off the proof, since $((\zeta\xi_0\xi_1... \xi_s)^p\zeta\xi^*_0, \zeta\xi^*_0(\xi^*_1... \xi^*_s\zeta^*_0)^p)$ contains diamonds that collaps in Σ_A. □

We write $\upsilon(\Sigma_{A(l)})$ for the subshift over the alphabet $\upsilon(\mathbb{C}_k)$ with the transition matrix induced by υ, i.e. for $\zeta,\zeta' \in \upsilon(\mathbb{C}_k)$ we set $\zeta \to \zeta'$ whenever $\chi \to \chi'$, where $(\zeta,\zeta') \in \upsilon(\chi)\times\upsilon(\chi')$ for some $\chi,\chi' \in \mathbb{C}_k$.

Theorem 10:

(i) If $\upsilon(\Sigma_{A(k)}) = \Sigma_{A(l)}$ then $\Omega_k \cong \Omega_l$ and

(ii) if $\upsilon(\Sigma_{A(k)}) \ne \Sigma_{A(l)}$ for all Σ_k with $\aleph(k) = \aleph(l)+1$, then transitive points in

$\Sigma_{A(l)}$ have trivial equivalence classes.

Proof. Clearly, $\upsilon(\mathbb{C}_k) = \mathbb{C}_l$; and it follows by lemma 9 that the transition matrix induced by υ on $\mathbb{C}_l$ coincides with $A(l)$.

(ii) Suppose ξ is a transitive point in $\Sigma_{A(l)}$ with non-trivial equivalence class. By lemma 8 it follows that $\xi_i \neq \zeta_i$ for any $\zeta \in \langle\xi\rangle$. Set $\xi_i = \{\xi_i^0,\ldots,\xi_i^m\}$, $\zeta_i = \{\zeta_i^0,\ldots,\zeta_i^m\}$, where $m = \eth(l)$, and let $j\colon \mathbb{Z} \to \{0,\ldots,m\}$ be indices chosen so that $\zeta_i^{j(i)} \neq \xi_i^s$ for $s = 0,\ldots,m$ and $i \in \mathbb{Z}$; then $\xi^* = \{\xi_i^0,\ldots,\xi_i^m,\zeta_i^{j(i)}\}$ is an element in $\mathbb{C}$ and lies necessarily in some $\mathbb{C}_k$, where $\eth(k) = \eth(l)+1$. In particular, since ξ is transitive it realises every possible transition, therefore in $\mathbb{C}$ exists a ξ^* of one higher dimension so that $\upsilon(\xi^*) = \xi$. But this means that $\upsilon(\Sigma_{A(k)}) = \Sigma_{A(l)}$ and by the first part of the theorem $\Omega_k \cong \Omega_l$. $\square$

We call a subshift $\Sigma_{A(l)}$ or subalphabet $\mathbb{C}_l$ reduced if $\mathbb{C}_l$ satisfies the condition in part (ii) of the theorem. We can always find a chain $l[0], l[1],\ldots, l[p]$ of some length p, so that

(i) $\eth(l[q+1]) = \eth(l[q])+1$ for $q = 0,\ldots,p-1$,

(ii) $\Sigma_{A(l[q+1])}/\approx\ \cong \Sigma_{A(l[q])}/\approx$ for $q = 0,\ldots,p-1$ and

(iii) transitive points in $\Sigma_{A(l[p])}$ have trivial equivalence classes.

By virtue of theorem 10 there exists always a reduced subshift of finite type whose quotient is isomorphic to $\Omega = \Sigma_A/\approx$. For the rest of this paper we shall assume Σ_A is already reduced, i.e. transitive points in Σ_A have trivial equivalence classes.

5. The product structure on Ω.

A local product structure is a map $[\cdot,\cdot]\colon \Omega\times\Omega \to \Omega$ defined in a neighbourhood of the diagonal of $\Omega\times\Omega$ and which has the properties

(i) $[x,x] = x$, $[[x,y],z] = [x,[y,z]] = [x,z]$, $[\sigma x,\sigma y] = \sigma[x,y]$; whenever these expressions are defined;

(ii) there exist $\delta > 0$, $\lambda \in (0,1)$ so that

(α) if $d(y_i,x) < \delta$ and $[y_i,x] = y_i$, $i = 1,2$, then $d(\sigma^n y_1,\sigma^n y_2) \leq \lambda^n d(y_1,y_2)$ for $n > 0$;

(β) if $d(x,z_i) < \delta$ and $[x,z_i] = z_i$, $i = 1,2$, then $d(\sigma^n z_1,\sigma^n z_2) \leq \lambda^{|n|} d(z_1,z_2)$ for $n < 0$;

The point $[x,y]$ lies in the stable direction of x and in the unstable direction of y. See also [6] p. 125 ff.

Recall that Σ_A is assumed to be mixing, i.e. $A^n > 0$ for large n. We define the one-sided shift space

$$\Sigma_A^+ = \{x \in \prod_{0\ldots+\infty} \mathfrak{A}\colon A[x_i,x_{i+1}] = 1 \;\forall\; i \in \mathbf{Z}^+\},$$

and similarly

$$\Sigma_A^- = \{x \in \prod_{-\infty\ldots 0} \mathfrak{A}\colon A[x_i,x_{i+1}] = 1 \;\forall\; i \in \mathbf{Z}^-\}.$$

Denote by $\mathfrak{N}^+(a)$ the set of all words (cylinder) $x_0x_1\ldots \in \Sigma_A^+$ that begin with $x_0 = a$. Similarly $\mathfrak{N}^-(a) = \{\ldots x_{-1}x_0 \in \Sigma_A^-\colon\; x_0 = a\}$. Two sequences $x_0x_1\ldots$, $y_0y_1\ldots$ in Σ_A^+ are related, write $x_0x_1\ldots \sim y_0y_1\ldots$, whenever $x_i \sim y_i$ for all $i = 0,1,\ldots$; and similarly for Σ_A^-. Given $\zeta \in \mathfrak{N}^\pm(a)$ we then put

$$\mathfrak{S}^\pm(\zeta) = \{\xi \in \Sigma_A^\pm\colon\; \xi \sim \zeta\},$$

for the set of all half - infinite words that are related to ζ. Denote by π_i the projection onto the i-th coordinate; in particular

$$\pi_0\mathfrak{N}^-(\zeta) = \{y_0\colon\; \ldots y_{-1}y_0 \in \mathfrak{S}^-(\zeta)\},$$

$$\pi_0\mathfrak{N}^+(\zeta) = \{y_0\colon\; y_0y_1\ldots \in \mathfrak{S}^+(\zeta)\}.$$

In addition to $\mathfrak{C}$ and Σ_C let us introduce corresponding one-sided objects $\mathfrak{C}^+$, $\mathfrak{C}^-$, Σ_C^+, Σ_C^-. We take the collection of sub alphabets $\{\mathfrak{A}_n^k\colon\; k,n\}$ which was introcduced in section 2, and prune away all those elements that cannot extended infinitly in forward direction. This defines $\mathfrak{C}^+$. Similarly $\mathfrak{C}^-$ is defined as $\{\mathfrak{A}_n^k\colon\; k,n\}$ less those $\mathfrak{A}_n^k$ that cannot extended infinitly into backward direction. Furthermore we define transition matrices C^+, C^- (and similarly $A^+(i)$, $A^-(i)$) in the same fashion as done in section 2, and call the associated one sided shift spaces Σ_C^+, Σ_C^-. For convenience we agree on the notation '±' whenever we would like to write a formula with either of them.

Observe

(i) $\mathfrak{C}^+ \cap \mathfrak{C}^- = \mathfrak{C}$,

(ii) the transitive subspaces of Σ_C^+, Σ_C^- are the same as those of Σ_C.

(iii) given $\mathfrak{u}_0\mathfrak{u}_1\mathfrak{u}_2 \ldots \in \Sigma_C^+$, then the cut off $\mathfrak{u}_{\beta'}\mathfrak{u}_{\beta'+1}\mathfrak{u}_{\beta'+2}\ldots$ is a half - infinite sequence that occours in Σ_C, $0 \le \beta' \le \beta$. The same applies to Σ_C^-: If $\ldots\mathfrak{u}_{-2}\mathfrak{u}_{-1}\mathfrak{u}_0 \in \Sigma_C^+$, then $\ldots\mathfrak{u}_{\beta''-2}\mathfrak{u}_{\beta''-1}\mathfrak{u}_{\beta''}$ is a sequence in Σ_C, $-\beta \le \beta'' \le 0$.

Analoguously to the two-sided case, there is a partial ordering on the elements in $\Sigma_C^\pm$. Given $\zeta \in \Sigma_C^\pm$, then $\mathfrak{S}^\pm(\zeta)$ consists of the maximal elements in $\Sigma_C^\pm$ that contain ζ. Define $\mathfrak{M}^+$ as all $\mathfrak{u}_0$ so that there exists a half infinite word $\mathfrak{u}_0\mathfrak{u}_1\mathfrak{u}_2\ldots$ which is maximal in Σ_C^+. Similarly $\mathfrak{M}^- = \{\mathfrak{u}_0 : \exists \ldots\mathfrak{u}_{-2}\mathfrak{u}_{-1}\mathfrak{u}_0$ maximal in $\Sigma_C^-\}$. If we put $\mathfrak{M}$ for the sub alphabet of $\mathfrak{C}$ that containes the symbols out of which maximal strings of Σ_C are composed, we have in particular $\mathfrak{M} = \{\mathfrak{u} \cap \mathfrak{v} : (\mathfrak{u}, \mathfrak{v}) \in \mathfrak{M}^- \times \mathfrak{M}^+\} \backslash \{\emptyset\}$. For $a \in \mathfrak{A}$ set $\mathfrak{M}^\pm(a) = \{\mathfrak{u} \in \mathfrak{M}^\pm : a \in \mathfrak{u}\}$ and define maps $\mu^+, \mu^- : A \to 2^{\mathfrak{A}}$ by setting

$$\mu^\pm(a) = \bigcap \{\mathfrak{u} \in \mathfrak{M}^\pm(a)\}.$$

On has to think of points in, e.g. $\mu^+(a)$ as adhering to a under future continuation. For a subset $\mathfrak{G} \subset \mathfrak{A}$ we set generally $\mu^\pm(\mathfrak{G}) = \bigcup_{g \in \mathfrak{G}} \mu^\pm(g)$. Note that $\mathfrak{G} \subset \mu^\pm(\mathfrak{G})$. It takes finitly many steps to construct $\mathfrak{C}^\pm$, $\mathfrak{M}^\pm$ (by checking strings which length is at most $\beta = |2^{\mathfrak{A}}|$).

Lemma 11: $\mu^- \circ \mu^- = \mu^-$ and $\mu^+ \circ \mu^+ = \mu^+$.

Proof. We shall do the proof only for μ^-, it works exactly the same for μ^+. Given some $a \in \mathfrak{A}$, from the definition of μ^- it follows that for $b \in \mu^-(a)$ and for every string $\zeta \in \mathfrak{R}^-(a)$, the intersection $\mathfrak{S}^-(\zeta) \cap \mathfrak{R}^-(b)$ is not empty. Choose $c \in \mu^-(\mu^-(a)) = \bigcup_{b \in \mu^-(a)} \mu^-(b)$, then there exists a $b \in \mu^-(a)$ so that for every choice of $\xi \in \mathfrak{R}^-(b)$ the intersection $\mathfrak{S}^-(\xi) \cap \mathfrak{R}^-(c)$ is not empty. In particular we choose $\xi \in \mathfrak{S}^-(\zeta) \cap \mathfrak{R}^-(b)$. In other words, given $\zeta \in \mathfrak{R}^-(a)$ then $\mathfrak{S}^-(\zeta) \cap \mathfrak{R}^-(c) \neq \emptyset$ which by definition of μ^- proves that c lies in $\mu^-(a)$. Therefore $\mu^-(\mu^-(a)) = \mu^-(a)$. $\square$

We gather from the last lemma

$$b \in \mu^{\pm}(a) \leftrightarrow \mu^{\pm}(b) = \mu^{\pm}(a).$$

Thus, μ^-, μ^+, each decomposes $\mathfrak{S}$ into disjoint subsets. Furthermore, suppose $\mu^-(a)$ *consists of more than one single point. If* $b \in \mu^-(a)\setminus\{a\}$ *then necessarily* $\mathfrak{S} \le \{a,b\}$. (Using the convention: For $w \in \mathfrak{C}$, we write $w \le \mathfrak{S}$ ($\mathfrak{S} \le w$) if there exists $c \in \mathfrak{S}$ so that $w \le c$ ($c \le w$). Since Σ_A is topologically mixing this property does not depend on the particular choice of c.) To clarify the last statement observe that for any $\varsigma \in \mathfrak{Z}^-(a)$ the intersection $\mathfrak{F}^-(\varsigma) \cap \mathfrak{Z}^-(b)$ is non empty. Since transitive points are supposed to have trivial equivalence classes, there must necessarily be at least one transition $x_{m-1} \to \{x_m,y_m\} \in C$, $m \le 0$, for some $\ldots x_{-1}x_0 \in \mathfrak{Z}^-(a)$, and where $\ldots y_{-1}y_0 \in \mathfrak{F}^-(\ldots x_{-1}x_0) \cap \mathfrak{Z}^-(b)$. Analoguously, if $|\mu^+(a)| \ge 2$, then $\{a,b\} \le \mathfrak{S}$, for every $b \in \mu^+(a)\setminus\{a\}$. Since no diamonds in Σ_A collapse, it is therefore clear that $\mu^-(a)$ and $\mu^+(b)$, $a,b \in \mathfrak{S}$, intersect in at most one point.

Given $w = \{a,b\} \in \mathfrak{C}$, then we define relative cylinders $\mathfrak{Z}^+(a|b)$, $\mathfrak{Z}^-(a|b)$ as follows:

$$\mathfrak{Z}^+(a|b) = \{x_0x_1\ldots \in \mathfrak{Z}^+(a): \text{ there exists } y_0y_1\ldots \in \mathfrak{F}^+(x_0x_1\ldots) \cap \mathfrak{Z}^+(a)$$
$$\text{satisfying } \{x_0,y_0\} < \ldots < \{x_s,y_s\} \le \{x_s,y_s\} \text{ for some } s \ge 0\}.$$

In the same manner $\mathfrak{Z}^-(a|b)$ is defined. Obviously on has the inclusion $\mathfrak{Z}^{\pm}(a|b) \subset \mathfrak{Z}^{\pm}(a)$. Some more notation: For $\varsigma \in \mathfrak{Z}^{\pm}(a|b)$ we set $\mathfrak{F}^{\pm}(\varsigma|b) = \mathfrak{F}^{\pm}(\varsigma) \cap \mathfrak{Z}^{\pm}(a|b)$ and define $\mathfrak{M}^{\pm}(a|b) = \{\pi_0\mathfrak{F}^{\pm}(\varsigma|b): \varsigma \in \mathfrak{Z}^{\pm}(a|b)\}$. For $a \in \mathfrak{S}$ define the predecessor and sucessor sets as indicated

$$\mathfrak{P}(a) = \{c \in \mathfrak{S}: A[c,a] = 1\},$$
$$\mathfrak{f}(a) = \{c \in \mathfrak{S}: A[a,c] = 1\}.$$

For $\mathfrak{u} \subset \mathfrak{S}$ a subset, we write $\mathfrak{P}(\mathfrak{u}) = \bigcup_{c\in\mathfrak{u}} \mathfrak{P}(c)$, and similarly $\mathfrak{f}(\mathfrak{u})$. We now indroduce functions τ^-,τ^+ that are defined on pairs of related symbols and map into $\mathfrak{S}$. More precisely:

$$\tau^+(a|b) = \bigcap \{\mathfrak{P}(u): \mathfrak{u} \in \mathfrak{M}^+(a|b)\},$$
$$\tau^-(a|b) = \bigcap \{\mathfrak{f}(u): \mathfrak{u} \in \mathfrak{M}^-(a|b)\},$$

for $\{a,b\} \in \bigcup \{\mathfrak{C}_l: \mathfrak{d}(l)=2\}$. Finally, for $w = \{a,b\} \in \mathfrak{C}$ define $\mathfrak{Z}^-(w) = \bigcup \{\mathfrak{u} \in \mathfrak{C}:$

$\mathfrak{w} \to w\}$, and in the same way $\mathfrak{K}^+(w) = \bigcup \{\mathfrak{w} \in \mathfrak{C}: \mathfrak{w} \to w\}$. Later on we shall write $\mathfrak{u} \backslash \mathfrak{K}^{\pm}(w)$ instead of $\mathfrak{u} \cap \mathfrak{S} \backslash \mathfrak{K}^{\pm}(w)$, $\mathfrak{u}$ a subset of $\mathfrak{S}$.

We call $w = \{a,b\} \in \mathfrak{C}$ **isolated** if

(i) neither $\mathfrak{S} \leq w$ nor $w \leq \mathfrak{S}$ is satisfied;

(ii) at least one of the intersections $\{\mathfrak{u} \cap \mathfrak{v}: (\mathfrak{u},\mathfrak{v}) \in \mathfrak{M}^{\mp}(a|b) \times \mathfrak{M}^{\pm}(b|a)\}$ is empty.

The next lemma is a immediate consequence of this definition.

Lemma 12: There is no local product structure on Ω if $\mathfrak{C}$ has isolated elements.

Proof: *Given some $\delta > 0$, and denote by $U_\delta(x)$ the δ-ball around $x \in \Sigma_A$.* We shall construct a sequence of points $(x[q],y[q]) \in \Sigma_A \times \Sigma_A$ that converges to the diagonal in $\Sigma_A/\approx \times \Sigma_A/\approx$ and has the property that the local stable direction through $x[q]$ and the local unstable direction through $y[q]$ have empty intersection. If we denote by $U_\delta(x)$ the δ-ball around $x \in \Sigma_A$, where δ desribes the size of the local stable and unstable directions, this means in particular that on cannot find a positive ε so that $d(x,y) \leq \varepsilon$ entails necessarily that $W^s(y) \cap U_\delta(y)$ and $W^u(x) \cap U_\delta(x)$ have a non empty intersection.

We shall construct a sequence of points $(x[q],y[q]) \in \Sigma_A \times \Sigma_A$ so that $d(x[q],y[q]) \leq 2^{[qr/2\alpha]}$ (for some $r \geq 1$) and so that the intersection $W^s(y[q],qr) \cap W^u(x[q],qr)$ is empty. This contradicts the continuity property of a local product structure. Let $w = \{a,b\}$ be an isolated element, then in particular there exist $(\mathfrak{u},\mathfrak{v}) \in \mathfrak{M}^+(a|b) \times \mathfrak{M}^-(b|a)$ satisfying $\mathfrak{u} \cap \mathfrak{v} = \emptyset$. Let $x_0x_1\ldots \in \Sigma_A{}^+$, $\ldots y_{-1}y_0 \in \Sigma_A{}^-$, $(x_0,y_0) = (a,b)$, be sequences, so that

$$\pi_0 \mathfrak{B}^-(x_0x_1\ldots) = \mathfrak{u},$$

$$\pi_0 \mathfrak{B}^+(\ldots y_{-1}y_0) = \mathfrak{v}.$$

By definition of $\mathfrak{M}^-(b|a)$ there exists a Σ_A-word $x_s\ldots x_{-1}$, $x_{-1} \to x_0$, related to $y_s\ldots y_1$, $s \leq 0$, so that $\{x_s,y_s\} \leq \{x_s,y_s\}$. Let $\{x_s,y_s\}\{x_{s-1},y_{s-1}\}\ldots\{x_s,y_s\}$ be a Σ_C-loop of length $l+1 \geq 1$ and put $(\nu',\tau') = (x_{s-l}\ldots x_s, y_{s-l}\ldots y_s)$. Similarly, there exists a Σ_A-word $y_1 \dots y_t$, $y_0 \to y_1$, related to $x_1 \dots x_t$, $t \geq 0$ so that $\{x_t, y_t\} \leq \{x_t, y_t\}$. Analoguously, let $\{x_t, y_t\} \dots \{x_{t+k}, y_{t+k}\}\{x_t, y_t\}$ be a Σ_C-loop and put $(\nu'', \tau'') = (x_t \dots x_{t+k}, y_t \dots y_{t+k})$. For $q \geq 1$ put

$$x[q] = \qquad \dots \nu'^q x_s \dots \mathbf{x_0} x_1 \dots x_t \nu''^q x_{t+1} x_{t+2} \dots,$$

$$y[q] = \dots y_{s-2} y_{s-1} \tau'^q y_s \dots \mathbf{y_0} y_1 \dots y_t \tau''^q \dots,$$

where bolds characters denote the zero position and the dots to the left of ν' and to the right of τ'' denote anything that makes $x[q], y[q]$ to one - sided transitive points in Σ_A. Set

$$\mathfrak{u}^q = \pi_0 \mathfrak{S}^+(x_0 x_1 \dots x_t \nu''^q x_{t+1} x_{t+2} \dots) \subset \mathfrak{u},$$

$$\mathfrak{v}^q = \pi_0 \mathfrak{S}^-(\dots y_{s-2} y_{s-1} \tau'^q y_s \dots y_{-1} y_0) \subset \mathfrak{v},$$

and we have necessarily $\mathfrak{u}^q \cap v^q = \emptyset$, $q \geq 1$. By construction $x[q]$ and $y[q]$ are not equialent, and if there were a $z \in W^s(x[q], \beta) \cap W^u(y[q], \beta)$ then the zero's coordinate in paricular would have the property

$$z_0 \in \mathfrak{u}^q \cap \mathfrak{v}^q,$$

which is assumed to be empty. On the other hand we have $d(x[q], y[q]) \leq 2^{[qr/2\alpha]}$, where $r = \min\{k, l\}$. □

Lemma 13: Suppose $w = \{a,b\}$ is an element in $\mathfrak{C}$ that satisfies either $\mathfrak{A} \leq w$ or $w \leq \mathfrak{A}$. Then

(i) if $\mathfrak{A} \leq w$, if $c \in \mathfrak{P}(a) \cap \mathfrak{K}^-(w)$ then $\mu^-(c)$ and $\tau^+(b|a)$ are disjoint;

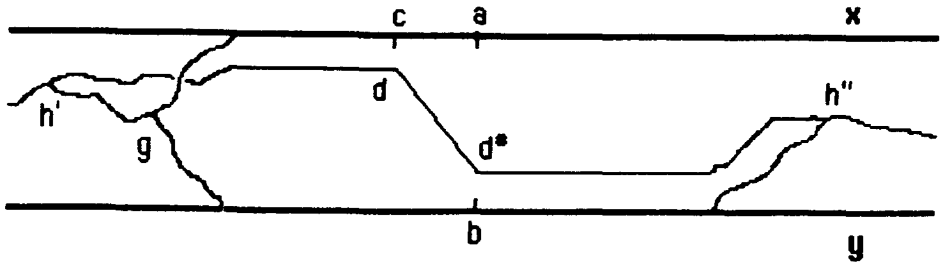
(ii) if $w \leq \mathfrak{A}$, if $c \in \mathfrak{f}(a) \cap \mathfrak{K}^+(w)$ then $\mu^+(c)$ and $\tau^-(b|a)$ are disjoint.

Proof. We shall elaborate only the first part $\mathfrak{A} \leq w$, the second case (ii) works exactly the same. The proof is by contradiction, we assume the intersection were not empty, we shall then construct a collapsing diamond. Let $w_0 = \{a,b\} \in \mathfrak{C}$ be a such, that for some $c \in \mathfrak{P}(a) \cap \mathfrak{K}^-(w_0)$ the intersection $\mu^+(c) \cap \tau^-(b|a)$ is not empty. Choose $d \in \mu^+(c) \cap \tau^-(b|a)$. Clearly, one can find a Σ_C-word $w_k w_{k+1} \dots w_0 w_1 \dots w_l$, for some $k \leq 0 \leq l$, $|w_i| \leq 2$ for $i = k, \dots, l$, and $w_k \in \mathfrak{C}_{k'}$, $w_l \in \mathfrak{C}_{l'}$ for some $\mathfrak{C}_{k'}, \mathfrak{C}_{l'}$ so that $\Sigma_{A(k')}, \Sigma_{A(l')}$ are both transitive. We may assume that there are exactly two Σ_A-words τ, ν running through $w_k \dots w_l$

(chosen so that $(\tau_0,\nu_0) = (a,b)$). (If there were more than two words, then two - element w that can be replaced by one - element w restricting the number of Σ_A-words running through the transformed Σ_C-word. This procedure can be repeated until there are exactly two strings running through the eventual $w_k \dots w_l$.) Let $\{x'_0,y'_0\}\dots\{x'_{p'-1},y'_{p'-1}\}\ \{x'_0,y'_0\}$, be a $\Sigma_{A(k')}$-loop satisfying $\{x'_{p'-1},y'_{p'-1}\} \to w_k$. Set $(\tau',\nu') = (x'_0x'_1\dots x'_{p'-1},\ y'_0y'_1\dots y'_{p'-1})$. Similarly there exists a $\Sigma_{A(l')}$-loop $\{x''_0,y''_0\}\dots\ \{x''_{p''-1},y''_{p''-1}\}\ \{x''_0,y''_0\}$ that satisfies $w_k \to \{x''_0,y''_0\}$. Set $(\tau'',\nu'') = (x''_0x''_1\dots x''_{p''-1},\ y''_0y''_1\dots y''_{p''-1})$ and define $(x,y) = (\tau'^{\infty}\tau\tau''^{\infty},\ \nu'^{\infty}\nu\nu''^{\infty})$.

Now shall proceed to construct a collapsing diamond. Set $q = \max([\beta/p'] + 1, [\beta/p''] + 1)$. By assumption $\mathfrak{G} \le w_0$, and clearly the same holds true for w_k, i.e. $\mathfrak{G} \le w_k$. Hence there exist related Σ_A-words (of the same length, naturally) γ,δ and an element $g \in \mathfrak{G}$, so that $g\gamma\nu' \sim g\delta\tau'$. Let $\zeta' \in \Sigma_A^-$ chosen so that $\zeta'g\gamma\nu'$ again lies in Σ_A^-. Since $d \in \mu^-(c)$, $c \in \mathfrak{F}(a)$, for every choice of ζ' there exists at least one $\xi' \in \Sigma_A^-$ satisfying $\xi'd \in \mathfrak{F}^-(\zeta'g\gamma\tau'^q\tau^-)$, where $\tau^- = \tau_k\tau_{k+1}\dots\tau_{-1}$. (The inset τ'^q streches related section and enables us later on to apply lemma 6.) We may assume that ζ' is negatively transitive in Σ_A. In particular this entails that $\xi'd$ and $\zeta'g\gamma\tau'^q\tau^-$ agree for large enough negative indices. Let h' be the first symbol they have in common and set ζ' $\chi'h'\zeta'^*$, $\xi' = \chi'h'\xi'^*$ for some $\chi' \in \Sigma_A^-$ and ζ'^*,ξ'^* finite Σ_A-words.

For positive coordinates we do a similar construction as follows. Set $\nu^+ = \nu_0\nu_1\dots\nu_l$ and pick a positively Σ_A-transitive continuation $\nu^+\nu''^q\zeta'' \in \Sigma_A^+$, $\zeta'' \in \Sigma_A^+$, satisfying $\pi_i\mathfrak{F}^+(\nu^+\nu''^q\zeta'') = \{(\nu^+\nu''^q\zeta'')_i\}$ for large enough $i \in \mathbb{N}$. Since $w = \pi_0\mathfrak{F}^+(\nu^+\nu''^q\zeta''|a) \in M^+(b|a)$, by assumption there exists a transition $d \to d^* \in w$. Let $d^*\xi''$ be an element in $\mathfrak{F}^+(\nu^+\nu''^q\zeta''|a)$, then there one finds $h'' \in \mathfrak{G}$, $\chi'' \in \Sigma_A^+$ and finite Σ_A-words ζ''^*,ξ''^* so that $\nu^+\nu''^q\zeta'' = \nu^+\nu''^q\zeta''^*h''\chi''$ and $d^*\xi'' = d^*\xi''^*h''\chi''$. We have the following situation, (where $(x,y) = (\tau'^{\infty}\tau\tau''^{\infty},\ \nu'^{\infty}\nu\nu''^{\infty})$)



One uses lemma 6 (with $k = \beta$) on the three points $(x,y,w) = (\zeta' g\gamma\tau'^{q}\tau\tau''^{\infty}, \zeta' g\delta\nu'^{q}\nu\nu''^{q}\zeta'', \xi' dd^{*}\xi'')$ and concludes that $\zeta' g\nu'^{q}\nu^{-}$ is related to $\xi' d$. Hence $(h'\zeta'^{*} g\delta\nu'^{q}\nu\nu''^{q}\zeta''^{*}h'', h'\xi'^{*} dd^{*}\xi''^{*}h'')$ is a diamond that collapses under $\sim$. This proves the lemma. $\square$

The next propositition will discuss a sufficient and necessary conditions for the existence of a continuous local product structure on $\Sigma_A/\approx$.

Proposition 14: There is a local product structure on Ω exactly if

(i) for all $w = \{a,b\} \in \mathfrak{C}$ that satisfy either $\mathfrak{S} \le w$ or $w \le \mathfrak{S}$ we have

(A) if $\mathfrak{S} \le w$, $(\mu^{-}\mathfrak{W}(a)\backslash\mathfrak{A}^{-}(w), \mu^{-}\mathfrak{W}(b)\backslash\mathfrak{A}^{-}(w)) \subset \mu^{-}\tau^{+}(b|a)\times\mu^{-}\tau^{+}(a|b)$;

(B) if $w \le \mathfrak{S}$, $(\mu^{+}\mathfrak{F}(a)\backslash\mathfrak{A}^{+}(w), \mu^{+}\mathfrak{F}(b)\backslash\mathfrak{A}^{+}(w)) \subset \mu^{+}\tau^{-}(b|a)\times\mu^{+}\tau^{-}(a|b)$;

(ii) $\mathfrak{C}$ has no isolated elements.

Proof. (i) Let us first discuss the first half of part (i). Say:

(*) For every $c \in \mathfrak{W}(a)\backslash\mathfrak{A}^{-}(w)$ and $(v,w) \in \mathfrak{M}^{+}(b|a)\times\mathfrak{M}^{-}(c)$ the intersection $\mathfrak{W}(v) \cap w$ is not empty.

Consider $\mathfrak{W}(v) \cap w$ and take intersections over $(v,w) \in \mathfrak{M}^{+}(b|a)\times\mathfrak{M}^{-}(c)$, then (*) reads as $\tau^{+}(b|a) \cap \mu^{-}(c) \neq \emptyset$, for $c \in \mathfrak{W}(a)\backslash\mathfrak{A}^{-}(w)$ (since $\tau^{+}(b|a) = \bigcap \{\mathfrak{W}(u)\colon \mathfrak{u} \in \mathfrak{M}^{+}(b|a)\}$). Say $\tau^{+}(b|a) \cap \mu^{-}(c) = \mathfrak{u} \neq \emptyset$, in particular $\mathfrak{u} \subset \mu^{-}(c)$ and since $\mathfrak{u}$ is not empty we have necessarily by lemma 11 $\mu^{-}\circ\mu^{-}(c) = \mu^{-}(c) = \mu^{-}(\mathfrak{u})$. Hence

$$\mu^{-}(\tau^{+}(b|a) \cap \mu^{-}(c)) \subset \mu^{-}\tau^{+}(b|a) \cap \mu^{-}\circ\mu^{-}(c) = \mu^{-}\tau^{+}(b|a) \cap \mu^{-}(c) = \mu^{-}(c),$$

and therefore $\mu^{-}(c) \subset \mu^{-}\tau^{+}(b|a)$. This shows that (*) is equivalent to the first half of (i,A); the other cases are treated in the same way.

The rest of the first part of the proof consists of verifying the local product structure on Ω. Given $x,y \in \Sigma_A$ not too far apart, then we shall describe how $[x,y]$ is determined. Recall that $\beta = |2$[illegible]$|$, and pick $x,y \in \Sigma_A$, $d^*(x,y) \le \lambda^\beta$, and assume [illegible] $\le \{x_I,y_I\}$ for some $I \in [-\beta,\beta]$. We begin demonstrating the case of $[y,x]$. Let m be the infimum over all $p \le I$ for which there exist a Σ_C-word $\{x_p,y'_p\}\{x_{p+1},y'_{p+1}\}\cdots\{x_I,y'_I\}$, $y'_I = y_I$; if the minimum does not exist then we set $m = -\infty$. This procedure garanties that [illegible]$(x_m)\backslash$[illegible]$^-(\{x_m,y'_m\})$ is not empty. If $m = -\infty$ set $[y,x] = \ldots y'_{I-2}y'_{I-1}y_Iy_{I+1}\cdots$; and in case m is finite there exist by assumption half - infinite sequences

$$(\ldots x^\sim_{m-2}x^\sim_{m-1},\ y''_my''_{m+1}\cdots) \in [illegible]^-(\ldots x_{m-2}x_{m-1})\times[illegible]^+(y'_m\cdots y'_{I-1}y_Iy_{I+1}\cdots|x_m)$$

and a transition $x^\sim_{m-1} \to y''_m$. Then we define

$$[y,x] = \ldots x^\sim_{m-2}x^\sim_{m-1}y''_my''_{m+1}\cdots .$$

The restiction to [illegible]$\backslash$[illegible]$^-(\{a,b\})$ which was made in the statement of the proposition is a result of lemma 13. In particular, we see that $d([y,x],y) \le d(x,y)$ and $d([y,x],x) \le d(x,y)$. The first inequality is obvious, the second follows from lemma 6, since $[y,x]$ lies in the stable direction of y. To determin $[x,y]$ suppose [illegible] $\le \{x'_I,y_I\}$, $I \in [-\beta,\beta]$, and let $m' \le I$ be the infimum over all p such that there exists a Σ_C-word $\{x'_p,y_p\}\{x'_{p+1},y_{p+1}\}\cdots\{x'_I,y_I\}$, $x'_I = x_I$, and set in the case $|m'| < \infty$

$$[x,y] = \ldots y^\sim_{m'-2}y^\sim_{m'-1}x''_{m'}x''_{m'+1}\cdots ,$$

for sequences

$$(\ldots y^\sim_{m'-2}y^\sim_{m'-1},\ x''_{m'}x''_{m'+1}\cdots) \in [illegible]^-(\ldots y_{m'-2}y_{m'-1})\times[illegible]^+(x'_{m'}\cdots x'_{I-1}x_Ix_{I+1}\cdots|y_m);$$

and $[x,y] = \ldots x'_{I-2}x'_{I-1}x_Ix_{I+1}\cdots$ if $m' = -\infty$. The case $w \le$ [illegible] is dealt with in the very same manner, except that here $m, m' \ge I$ and we have a suitable transition from a 2-element symbol in [illegible] to a 1-element symbol.

We have to verify the identity $[[x,y],z] = [x,z]$ for points x, y, z have d^* distance $\lambda^{2\beta}$ from each other. Assume [illegible] $\le \{x_I,y_I\}$, $I \in [-\beta,\beta]$, then as was pointed out above

$$[x,y] = \ldots y^\sim_{m-2}y^\sim_{m-1}x''_mx''_{m+1}\cdots ,$$

for some $m \le I$. Consequently $d([x,y],x) \le d(x,y) \le \lambda^{2\beta}$, and suppose $\{x''_{I'},z_{I'}\} \le$ [illegible] for some $I' \in [\beta,2\beta]$. In the same way as above we find an integer $m' \ge I'$, sequences (if m' is finite)

$$(\ldots z'_{m'-2}z'_{m'-1},\ x^\sim_{m'}x^\sim_{m'+1}\cdots) \in [illegible]^-(\ldots y^\sim_{m-1}x''_m\cdots x''_{m'-1}|z_{m'})\times[illegible]^+(x''_{m'}x''_{m'+1}\cdots),$$

and a transition $z'_{m'-1} \to x^{\sim}_{m'}$, so that

$$[[x,y],z] = \ldots z'_{m'-2} z'_{m'-1} x^{\sim}_{m'} x^{\sim}_{m'+1} \cdots .$$

We have by construction $[[x,y],z]_i \sim [x,y]_i \ \forall \ i \geq -2\beta$, $[x,y]_i \sim x_i \ \forall \ i \geq -2\beta$, and therefore because of corollary 2 $[[x,y],z]_i \sim x_i \ \forall \ i \geq -2\beta+\alpha$. The same argument provides $[[x,y],z]_i \sim z_i \ \forall \ i \leq 2\beta-\alpha$. Since $\alpha \leq \beta$ we conclude $[[x,y],z] \approx [x,z]$. Next let us assume $\mathfrak{X} \leq \{x''_{\Gamma}, z_{\Gamma}\}$, $\Gamma \in [-2\beta,-\beta]$, and let $m' \leq \Gamma$. In particular $m' < m$, and if $m' = -\infty$, then $W^u(y,m) = W^u(z,m)$ and we are finished. In the case $-\infty < m'$, we have

$$[[x,y],z] = \ldots z^{\sim}_{m'-2} z^{\sim}_{m'-1} x'''_{m'} x'''_{m'+1} \cdots ,$$

for sequences

$$(\ldots z^{\sim}_{m'-2} z^{\sim}_{m'-1},\ x'''_{m'} x'''_{m'+1} \ldots) \in \mathfrak{W}^-(\ldots z_{m'-2} z_{m'-1}) \times \mathfrak{W}^+(y_{m'} \ldots y_{m-1} x''_m x''_{m+1} \ldots | z_m),$$

and on the other hand we observe

$$[x,z] = \ldots z^{\wedge}_{m''-2} z^{\wedge}_{m''-1} x''_{m''} x''_{m''+1} \cdots ,$$

where

$$(\ldots z^{\wedge}_{m''-2} z^{\wedge}_{m''-1},\ x''_{m''} x''_{m''+1} \ldots) \in \mathfrak{W}^-(\ldots z_{m''-2} z_{m''-1}) \times \mathfrak{W}^+(x_{m''} x_{m''+1} \ldots | z_{m''}).$$

By the same argument as above one finds $[x,z] \approx [[x,y],z]$.

(ii) Secondly, there is the possibility of strings with (possibly arbitrary) length greater that 2β made up of elements that satisfy neither $\mathfrak{X} \leq w$ nor $w \leq \mathfrak{X}$. Let $x,y \in \Sigma_A$, $d^*(x,y) \leq \lambda^{\beta}$. Since $\mathfrak{C}$ containes no isolated elements it follows that $\pi_0 \mathfrak{W}^-(\ldots y_{-1} y_0)$ and $\pi_0 \mathfrak{W}^+(x_0 x_1 \ldots)$ have non - empty intersection, and clearly $[x,y] = z$, where $(\ldots z_{-1} z_0,\ z_0 z_1 \ldots) \in \mathfrak{W}^-(\ldots y_{-1} y_0) \times \mathfrak{W}^+(x_0 x_1 \ldots)$. In exactly the same way one finds $[y,x]$. In the same way one verifies the formula $[[x,y],z] = [x,z]$ in the case that there are strings of length grater than 2β that consist entirely of elements that satisfy neither $\mathfrak{X} \leq w$ nor $w \leq \mathfrak{X}$. We shall not go into details.

(iii) We now proof the 'necessary' part. Parallel of the proof of lemma 12 we construct a sequence of points in $\Sigma_A \times \Sigma_A$ that converge in Ω to a periodic point. Suppose there exist $w = \{z_0, y^*_0\} \in \mathfrak{C}$, $x^*_0 \in \mathfrak{W}(z_0) \setminus \mathfrak{W}^+(w)$ and $(v,w) \in \mathfrak{W}^-(x^*_0) \times \mathfrak{W}^+(y^*_0 | z_0)$ so that $\mathfrak{W}(v)$ and w have empty intersection. Let

$(\ldots x^*_{-1}x^*_0, y^*_0y^*_1\ldots) \in \Sigma_A^- \times \Sigma_A^+$ be two half-infinite words satisfying $(\pi_0\mathfrak{F}^-(\ldots x^*_{-1}x^*_0), \pi_0\mathfrak{F}^+(y^*_0y^*_1\ldots|z_0)) = (v,w)$.

According to the definition of $\mathfrak{F}^\pm$ one finds a Σ_A-word $z_0z_1\ldots z_s$, $s \geq 0$, related to $y^*_0y^*_1\ldots y^*_s$, satisfying $\{z_0,y^*_0\} < \ldots < \{z_s,y^*_s\} \leq \{z_s,y^*_s\}$ (possibly $s = 0$). Pick a Σ_C-loop $\{x'_0,y'_0\}\ldots\{x'_{p-1},y'_{p-1}\}\{x'_0,y'_0\}$ satisfying $\{z_s,y^*_s\} \to \{x'_0,y'_0\}$, set $(\tau,\nu) = (x'_0x'_1\ldots x'_{p-1}, y'_0y'_1\ldots y'_{p-1})$ and define for $q \geq 1$ (remember: $x^*_0 \to z_0$)

$$x[q] = \ldots x^*_{-1}x^*_0z_0z_1\ldots\ldots z_s\tau^q\mathbf{x'_0}\ldots x'_{p-1}\tau^{q-1}\ldots,$$
$$y[q] = \ldots y^*_0y^*_1\ldots y^*_s\nu^q\mathbf{y'_0}\ldots y'_{p-1}\nu^{q-1}y^*_{s+1}y^*_{s+2}\ldots,$$

where dots to the left of τ and the right of ν denote anything in Σ_A, and the bold characters mark the zero position. From lemmas 4 and 5 it is clear that

$$\pi_0\mathfrak{F}^+(y^*_0y^*_1\ldots y^*_s\nu^q\mathbf{y'_0}\ldots y'_{p-1}\nu^{q-1}y^*_{s+1}y^*_{s+2}\ldots|z_0) \subset \pi_0\mathfrak{F}^+(y^*_1y^*_2\ldots|z_0).$$

Indeed $d^*(x[q],y[q]) \leq \lambda^{pq}$. If there were a continuous local product structure on $\Sigma_A/\approx$, there would exist a $\varepsilon > 0$ such that $d^*(x[q],y[q]) \leq \varepsilon$ implies $d^*([y[q],x[q]],x[q]) \leq \lambda$ and in particular $d^*(\sigma^{-s}[y[q],x[q]],\sigma^{-s}x[q]) \leq \lambda^{s+1}$, $s \geq 0$, since $[y[q],x[q]]$ lies on the unstable direction through $x[q]$. By construction $\sigma^{-pq}y[q] + \sigma^{-pq}x[q] = c \;\; \forall \;\; q \in \mathbb{N}$, and thus $d^*(\sigma^{-pq}[y[q],x[q]], \sigma^{-pq}x[q]) = 1 \geq \lambda$ while it is obvious that $x[q]$ and $y[q]$ converge to $\tau^\infty = x' \approx y' = \nu^\infty$. $\square$

6. The Zeta function on Ω.

Denote by $\mathfrak{N}(n)$ the number of periodic points in $\Omega = \Sigma_A/\approx$ with period n. The zeta function (See [7] p. 766) then is defined as

$$\zeta(t) = \exp\sum_{n\in\mathbb{N}} n^{-1}\mathfrak{N}(n)t^n,$$

for t a complex variable in some open set of $\mathbb{C}$. In particular we define for the individual subshifts $\Sigma_{A(i)}$:

$$\zeta_i(t) = \exp\sum_{n\in\mathbb{N}} n^{-1}\mathfrak{N}_i(n)t^n,$$

where $\mathfrak{N}_i(n)$ counts the the periodic points of period n in $\Sigma_{A(i)}$. In that case one has $\mathfrak{N}_i(n) = \text{trace } A(i)^n$ and the zeta functions are explicitly given by $\zeta_i(t) = [\det(\mathbf{1} - tA(i))]^{-1}$, which are analytic functions for $|t| < 1/h_i$, where h_i is the topological entropy of $\Sigma_{A(i)}$, i.e. the maximal (positive) eigenvalue of $A(i)$. Since the projection $\pi: \Sigma_A \to \Omega$ is finite to one, the topological entropy $h(\Sigma_A) = h_0 = h$ of Σ_A coincides with that of Ω. Furthermore one has $h_i \le h$ for all i. (cf. [2]). First let us prove an arithmetical lemma.

Lemma 15: Let $a[i] \ge 1$, $i = 1,2,\ldots,u$, be positive integers with at least one of them equal to 1, then

$$S = \sum_{1\le k\le u} \sum_{i[1],a[2],\ldots,i[k]} (-1)^{a[i[1]]+a[i[2]]+\ldots+a[i[k]]} = -1,$$

where in the second sum every index $i[j]$ appears at most once.

Proof. We prove the lemma by induction. Set $S = S(a[1],a[2],\ldots,a[u])$. For $u = 1$ it is clear that $S(a[1]) = -1$, since $(-1)^{a[1]} = -1$, $a[1] = 1$, is the only summand that appears in the sum. Suppose it is shown that $S(a[1],a[2],\ldots,a[u-1]) = -1$, then

$$S(a[1],a[2],\ldots,a[u]) = \sum_{1\le k\le u} \sum_{i[1],\ldots,i[k]} (-1)^{a[i[1]]+\ldots+a[i[k]]}$$

$$= \sum_{1\le k\le u-1} \Big\{\sum_{i[1],\ldots,i[k],i[j]\ne u} (-1)^{a[i[1]]+\ldots+a[i[k]]}$$

$$+ (-1)^{a[u]}\sum_{i[1],\ldots,i[k],i[j]\ne u} (-1)^{a[i[1]]+\ldots+a[i[k]]} + (-1)^{a[n]}\Big\}.$$

where the first sum consists of terms that do not contain $(-1)^{a[u]}$ and the second of these containing $(-1)^{a[u]}$. Hence

$$S(a[1],a[2],\ldots,a[u]) = S(a[1],a[2],\ldots,a[u-1]) + (-1)^{a[u]}[1 + S(a[1],a[2],\ldots,a[u-1])]$$

and therefore $S(a[1],a[2],\ldots,a[u]) = -1$. □

Theorem 16: $\log \zeta(t) = \sum_i (-1)^{\aleph(i)} \log \zeta_i(t)$ for $|t| < 1/h$, where $\aleph(i)$ is the dimension of $\mathfrak{C}_i$.

Proof. Pick a periodic point x in Σ_A, then $\langle x\rangle$ is a maximal element in Σ_C, has due to lemma 5 the same period n as x and lies by virtue of lemma 4 in a subshift of the form $\Sigma_{A(k)}$ for some k. This does not imply that all points in $\langle x\rangle$ have period n, instead of this $\langle x\rangle$ decomposes into subsets $\langle x\rangle^1$, $\langle x\rangle^2,\dots,\langle x\rangle^u$, $u \geq 1$, where $\langle x\rangle^i$ are sequences in Σ_C with the same period n as x, and are minimal in the sense that no $\langle x\rangle^1$ can be split into two nonempty $\langle x\rangle'$, $\langle x\rangle''$ which again have period n. One of the $\langle x\rangle^i$ is x itself. Denote by $a[i]$ the number of Σ_A-strings in $\langle x\rangle^i$, then $a[i]$ is exactly $d(i[j])$, where $i[j]$ is the index of the subshift $\Sigma_{A(i[j])}$ in which $\langle x\rangle^i$ lies. We now sum over all points in Σ_C that lie as sets in the equivalence class of x and have the same period; we count the Σ_C-points $\langle x\rangle^i$ with weight $(+1)$ if they have even dimension ($\partial = \gamma-1$) and with (-1) if they have odd dimension. Since at least one of the numbers $\gamma(i) = a[i]$ is one, we are lead to apply lemma 15, and obtain that the weighted sum is exactly 1.

There is one more complication. Up to now we neglected periodic points in Σ_C that lie in $\langle x\rangle$ but have a longer period than x. We order the points $y[j]$ of $\langle x\rangle$ with periods bigger than n, so that the period of $y[j]$ is not less than the one of $y[i]$ if $i < j$. For $\chi_1,\chi_2,\dots,\chi_v \in \Sigma_C$ we define $\bigvee(\chi_1,\chi_2,\dots,\chi_v)$ as the collection of all possible unions of χ_i that lie in Σ_C (not necessarily all unions lie in Σ_C). Set $U(0) = \bigvee(\langle x\rangle^1,\langle x\rangle^2,\dots,\langle x\rangle^u)$, and inductively $U(m) = \bigvee(\{y[m]\},U(m-1))$ (by construction $y[m] \notin \bigcup_{z\in U(m-1)} z$). Let us assume $\sum_{\chi\in U(m)} (-1)^{\partial(\chi)} = 1$, for some $m \geq 0$. We have proven this formula in the case $m = 0$, and indeed it holds true too for $m+1$, since

$$\sum_{\chi\in U(m)} (-1)^{\partial(\chi)} = \sum_{\chi\in U(m-1)} \left((-1)^{\partial(\chi)} + (-1)^{\partial(\chi)+1}\right) + 1 = 1.$$

The summand 1 on the left hand side comes from the single factor $1 = (-1)^{\partial(y[m])}$. We have therefore shown that on taking the weighted sum the counting of periodic points which have in Σ_A an equivalent point with smaller period cancelles out. Hence $\mathfrak{N}(n) = \sum_i (-1)^{\partial(i)} N_i(n)$ and the theorem follows (See also [5], theorem in §4.). □

In particular, this theorem proves rationality of the zeta function of Ω.

7. The subshift Σ_∞.

Let $\mathfrak{X}^* = 2^{\mathfrak{X}}$ be the power-set of $\mathfrak{X}$. We proceed to construct a nested sequence of shift-spaces of finite type and begin for $j \geq 2$ by defining j-dimensional transition matrices as follows:

$A_j(X_1,\ldots,X_j) = \{ 1$ if there are $x_k \in X_k$ for $1 \leq k \leq j$ which form a j-string, i.e. $A[x_k,x_{k+1}] = 1$ for all $k = 1,\ldots,j-1$;

$\{ 0$ otherwise (including the case in which one of the X_k is the empty set),

where $X_1,\ldots,X_j \in \mathfrak{X}^*$. The A_j define over the alphabet $\mathfrak{X}^*$ subshifts of type j which we call Σ_j. They form a nested sequence $\Sigma_j \supset \Sigma_{j+1}$ for $j \geq 2$. Define

$$\Sigma_\infty = \bigcap_{j\geq 2} \Sigma_j.$$

For $\lambda \in (0,1)$ there exists a natural metric on Σ_∞; for $x,y \in \Sigma_\infty$ we set $d(x,y) = \lambda^n$, where $n = \max\{m \in \mathbb{N}$: so that $x_i = y_i \ \forall \ |i| < m\}$. There is also a sequence of subshifts Σ_j' defined by restricting Σ_j to the points that have at least two one-element subsets of $\mathfrak{X}$ in all j-words. They act like the eye of a needle, and we have $\Sigma_j' \subset \Sigma_{j+1}'$ for $j \geq 2$. Hence Σ_∞ can be approximated as well from the inside: Σ_∞ = closure of $\bigcup_{j\geq 2} \Sigma_j'$. To see this, we observe that for every sequence $\ldots X_{-1}X_0X_1\ldots$ in Σ_∞ there exists (at least) one sequence $\ldots x_{-1}x_0x_1\ldots$ in Σ_A for which $x_i \subset X_i$, $i \in \mathbb{Z}$. Clearly, $\ldots X_{-1}X_0X_1\ldots$ can be approximated by points $\ldots X'_{-1}X'_0X'_1\ldots$ in Σ_{2j}', where we set $X'_i = \{x_i\}$ for $(i+j)$ mod $2j = 0$, and $X'_i = X_i$ otherwise. In general Σ_∞ will no longer be a subshift of finite type. The next lemma provides a criteriom which allows to classify all Σ_A for which the associated Σ_∞ is of finite type. We begin with a definition.

Definition 17: The transition matrix A has the <u>loop property</u> if for every loop of pairs $(v_1,w_1)\ldots(v_m,w_m)$ with

(i) $A[v_i,v_{i+1}]A[w_i,w_{i+1}] = 1$ for all $1 \leq i < m$,

(ii) $(v_1,w_1) = (v_m,w_m)$,

(iii) $(v_i,w_i) \neq (v_k,w_k)$ for all $i \neq k$ and $1 \le i,k < m$,

there are indices p,q with $1 \le p,q < m$, so that

$$A[v_p,w_{p+1}]A[w_q,v_{q+1}] = 1.$$

Lemma 18: The subshift Σ_∞ is of finite type ($\le |\mathfrak{A}|^2+1$) if and only if A has the loop property.

Proof. First the direction '⇒': It will be shown that $\Sigma_{j+1} = \Sigma_j$ for j large enough. Take a string $X_0...X_j$ where the X_i are elements in $\mathfrak{A}^*$, i.e. subsets of $\mathfrak{A}$. By assumption $X_0...X_{j-1}$ and $X_0...X_j$ are j-words in Σ_j. Hence there are j-strings $v_0...v_{j-1}$ and $w_1...w_j$ with $v_i \in X_i$ and $w_i \in X_i$ for $i = 0,...,j$. If $j \ge |\mathfrak{A}|^2+1$ there must be a pair (v_k,w_k) which appears twice. By the assumption there exists an index $p \in [0,j)$ so that $A[v_p,w_{p+1}] = 1$. Thus $v_0...v_p w_{p+1}...w_j$ is a $(j+1)$-word running through $X_0...X_j$.

The second part '⇐': Suppose Σ_∞ is of type n and that there is a loop of length j as in the statement of the lemma but having $A[v_i,w_{i+1}] = 0$ for all $i = 1,...,j-1$, then we will construct an $(n+1)$-string in Σ_∞ which satisfies the criterion on n-strings but fails for that on $(n+1)$-strings. Set V for the $(j-1)$-word $\{v_1\}...\{v_{j-1}\}$ and W for $\{w_k\}...\{w_{j-1}\}\{w_1\}...\{w_{k-1}\}$, where $k = n \bmod (j-1)$. Set furthermore X_i for the sets $\{v_i,w_i\} \in A^*$ containing two elements. Then the string

$$...VVVXX...XX_1...X_{k-1}WWW...,$$

where the word $X = X_1X_2...X_{j-1}$ is $[n/(j-1)]$-times repeated, lies in Σ_n but not any more in Σ_{n+1}. The case $A[w_i,v_{i+1}] = 0$ for $i = 1,...,j-1$ is excluded in the same way. □

The loop property implies mixing: Two elements $a,b \in \mathfrak{A}$ can be joined up by a string of length less than $|\mathfrak{A}|^2$. To see this, choose $x,y \in \Sigma_A$ two sequences with $x_0 = a$ and $y_n = b$, where $n = |\mathfrak{A}|^2$. The block $(x_0,y_0)...(x_n,y_n)$ contains a loop and thus there is a transition $x_p \to y_{p+1}$ for some p in $[0,n)$. The converse in general is not true as the following example demonstrates. Take the alphabet $\{1,2,3,4\}$ and define the transition matrix $A =$

$$\begin{vmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{vmatrix}.$$

The subshift Σ_A is mixing but has not the loop property as is to be seen at (1,2)(1,2), since that would require the transitions $1 \to 2$ and $2 \to 1$.

If the subhift Σ_∞ is of finite type the next lemma tells what the quotient $\Sigma_A/\approx$ looks like for any equivalence relation $\approx$.

Theorem 19: Suppose $\approx$ is an equivalence relation on Σ_A, and Σ_∞ is of finite type, n say. Then $\Sigma_A/\approx$ is isomorphic to a subshift of finite type.

Proof. As before $\mathbf{S}^*$ denotes the power set of $\mathbf{S}$. Define a map Φ from Σ_A into $\mathbf{S}^{*\mathbf{Z}}$ by

$$\Phi(y)_i = \{x_i\colon\ x \in \Sigma_A \text{ and } x \approx y\}.$$

This map commutes with the shift and has the property that $\Phi(y) = \Phi(z)$ exactly if $y \approx z$. Every Σ_A-word $y'_{-n}..y'_n$ which is related to $y_{-n}..y_n$ ($y'_i \sim y_i\ \forall\ |i| \leq n$) can be completed to a point $y' \in \Sigma_A$ which is equivalent to y, since by assumption (Σ_∞ is of type n) there are transitions $y_{-s} \to y'_{-s+1}$, $y'_{t-1} \to y_t$ for some $0 < s,t \leq n$. To determine $\Phi(y)_0$ it suffices therefore to know the components of y on the positions in $[-n,n]$ and we conclude that Φ is a $(2n+1)$-block map, that is Φ maps Σ_A continuously into the full $\mathbf{S}^*$-shift. Thus $\Sigma^* = \Phi(\Sigma_A)$ is isomorphic to $\Sigma_A/\approx$. Denote by $\mathbf{C}$ the set of all different subsets $\Phi(x)_0$ with x ranging over Σ_A. This is the alphabet for a new subshift, its transition matrix I is defined by

$$I(X,Y) = \begin{cases} 1 & \text{if there exist } x,y \in \Sigma_A \text{ such that } \Phi(x)_0 = X,\ \Phi(y)_1 = Y \text{ and } x \approx y; \\ 0 & \text{otherwise,} \end{cases}$$

for $X, Y \in C$. It remains to show that $\Sigma_I \cong \Sigma_A/\approx$:

(i) Choose $\xi \in \Sigma_I$, then there exists by construction a sequence x^k in Σ_A with $\Phi(x^k)_k = \xi_k$ and $x^k \approx x^{k+1}$ for $k \in \mathbf{Z}$. Hence $x^k \approx x^l$ for any $k,l \in \mathbf{Z}$, and $\Phi\colon \Sigma_A/\approx \to \Sigma_I$ is a surjection.

(ii) If $x,y \in \Sigma_A$ but $x \not\approx y$ then also $\Phi(x) \neq \Phi(y)$ which implies $\Phi(x)_k \neq \Phi(y)_k$

for an integer k.

This shows that the map $\Phi: \Sigma_A/\approx \to \Sigma_I$ is an isomorphism. □

References.

[1] R. Bowen; **On Axiom A Diffeomorphisms.** AMS Regional Conf. Proc., 35 (1978).

[2] R. Bowen - O. E. Lanford; Zeta functions of restrictions of the shift transformation. Proc. Symp. Pure Math. AMS, 14 (1970), pp 43 - 50.

[3] D. Fried; Metrique naturelles sur les espaces de Smale; C. R. Acad. Sc. Paris, Vol 297 (1983), pp 77 - 79.

[4] J. Kelly; **General Topology.** Van Nostrand.

[5] A. Manning; Axiom A diffeomorphisms have rational Zeta functions. LMS Bulletin, 3 (1971), pp 215 - 220.

[6] D. Ruelle; **Thermodynamic Formalism.** Addison Wesley, 1978.

[7] S. Smale; Differential Dynamical Systems. AMS Bulletin 73 (1967), pp 747 - 817.