Ark. Mat., **38** (2000), 281–317

Ergodic properties of fibered rational maps

Mattias Jonsson(1)

Abstract. We study the ergodic properties of fibered rational maps of the Riemann sphere. In particular we compute the topological entropy of such mappings and construct a measure of maximal relative entropy. The measure is shown to be the unique one with this property. We apply the results to selfmaps of ruled surfaces and to certain holomorphic mappings of the complex projective plane $\mathbf{P}^2$.

0. Introduction

Let X be a compact metric space, let $g: X \circlearrowleft$ be a continuous mapping, and let $\widehat{\mathbf{C}}$ denote the Riemann sphere. A *rational map of degree d fibered over g* is a continuous mapping $f: X \times \widehat{\mathbf{C}} \circlearrowleft$ of the form

$$f(x,z) = (g(x), Q_x(z)),$$

where Q_x is a rational function of degree d, depending continuously on $x \in X$. In this paper we will investigate the ergodic properties of fibered rational maps. For background on ergodic theory see e.g. [W].

In the special case when X is a point we recover the class of (non-fibered) rational maps of $\widehat{\mathbf{C}}$. The study of the ergodic properties of the latter mappings was initiated by Brolin [Br] (in the case of polynomials), and further developed by Lyubich, Freire, Lopez and Mañé. In particular they proved the following result.

Theorem A. ([L], [FLM], [M]) *Let $f: \widehat{\mathbf{C}} \circlearrowleft$ be a rational map of degree $d \geq 2$. Then the following holds*:

(i) *the topological entropy of f satisfies $h(f) = \log d$;*

(ii) *f has a unique measure μ of maximal entropy;*

(iii) *μ is mixing for f;*

(iv) *the support of μ is exactly the Julia set of f.*

(1) Supported by the Swedish Foundation for International Cooperation in Research and Higher Education (STINT).

In this paper we generalize Theorem A to the fibered setting (cf. Theorems 3.1, 4.2, 5.2 and 6.1). Let $\pi: X \times \widehat{\mathbf{C}} \to X$ be the natural projection.

Theorem B. *Let $f: X \times \widehat{\mathbf{C}} \circlearrowleft$ be a rational map of degree $d \geq 2$, fibered over $g: X \circlearrowleft$. Then the following holds:*

(i) *$h(f) = h(g) + \log d$;*

(ii) *if ν is an invariant probability measure for f, then the metric entropies of f and g are related by $h_\nu(f) \leq h_{\pi_*\nu}(g) + \log d$;*

(iii) *given an invariant probability measure μ' for g, there exists an invariant probability measure μ for f such that $\pi_*\mu = \mu'$ and $h_\mu(f) = h_{\mu'}(g) + \log d$; if $h_{\mu'}(g) < \infty$, then μ is unique with these properties;*

(iv) *the measure μ characterized in* (iii) *is ergodic (mixing) for f if μ' is ergodic (mixing) for g;*

(v) *if $x \in X$ and μ_x is the conditional measure of μ on $\{x\} \times \widehat{\mathbf{C}}$, then the support of μ_x is exactly the Julia set for the restriction of $\{f^n\}$ to $\{x\} \times \widehat{\mathbf{C}}$.*

In fact, we prove Theorem B in a more general setting. Namely, we replace the trivial $\widehat{\mathbf{C}}$-bundle $\pi: X \times \widehat{\mathbf{C}} \to X$ by a *$\widehat{\mathbf{C}}$-bundle* $\pi: Y \to X$. Roughly speaking, this means that Y is a compact metric space, that π is a continuous surjection, and that the fibers $Y_x := \pi^{-1}(x)$ are Riemann spheres which are fairly nicely packed together. A rational map fibered over $g: X \circlearrowleft$ is then a continuous mapping $f: Y \circlearrowleft$ such that π semiconjugates f to g, and such that the induced mappings $Y_x \to Y_{g(x)}$ between fibers are rational. See Section 1 for precise definitions.

The proof of Theorem B mainly follows Lyubich's proof of Theorem A. However, we use pluripotential theory to construct the measure of maximal entropy, as introduced by Hubbard and Papadopol [HP] and by Fornæss and Sibony [FS4]. Also, the proof of the uniqueness part in (iii) is substantially harder than in the non-fibered case. A different proof of the existence and uniqueness of the measure of maximal entropy has been given, independently, by Sumi [Su] for skew products generated by rational semigroups.

One motivation for studying fibered rational maps is that they can be used to understand the dynamics of certain holomorphic mappings in two complex dimensions. The first situation that we address is when Y is a *ruled surface*, i.e. a smooth projective variety with the structure of a $\mathbf{P}^1$-bundle $\pi: Y \to X$ over a compact Riemann surface X. By a result of Dabija [D], selfmaps of ruled surfaces can be viewed as fibered rational maps. Our techniques will therefore allow us to prove the following result (see Theorem 7.3).

Theorem C. *Let Y be a ruled surface over a Riemann surface X, and let f be a holomorphic mapping of Y that fibers over a holomorphic map $g: X \circlearrowleft$. Assume* *that the topological degrees* δ_f, δ_g *of* f *and* g *satisfy* $1<\delta_g<\delta_f$. *Then* $h(f)=\log\delta_f$, *and* f *has a unique measure of maximal entropy.*

For holomorphic mappings of complex projective space $\mathbf{P}^k$, $k\geq 2$, Fornæss and Sibony proved that there exists a natural measure of maximal entropy. The question of uniqueness is open in general, but we will prove the following result (see Theorem 7.4).

Theorem D. *Let* f *be a holomorphic selfmap of* $\mathbf{P}^2$ *of degree* $d\geq 2$ *that preserves the family of lines passing through a given point* O *in* $\mathbf{P}^2$. *Then* f *has a unique measure of maximal entropy.*

The last theorem applies in particular to *polynomial skew products* on $\mathbf{C}^2$. These are mappings of the form $f(z,w)=(p(z),q(z,w))$, where p and q are polynomials of degree $d\geq 2$, and q has nonzero w^d-term. Polynomial skew products on $\mathbf{C}^2$ were previously studied by Heinemann [H1], [H2], [H3], by the author [J], and by Sester [S1], [S2] (in a slightly different situation).

Fibered rational maps can be viewed as non-autonomous (or random) dynamical systems on $\widehat{\mathbf{C}}$. Indeed, if $f: X\times\widehat{\mathbf{C}}\circlearrowleft$ is a fibered rational map, then the restriction of f^n to $\{x\}\times\widehat{\mathbf{C}}$ defines a non-autonomous system on $\widehat{\mathbf{C}}$, where the time n-map is given by

$$z\longmapsto Q_{g^{n-1}(x)}\circ\ldots\circ Q_{g(x)}\circ Q_x(z).$$

Conversely, let $(Q_i)_{i\geq 0}$ be an equicontinuous sequence of rational maps of $\widehat{\mathbf{C}}$ of degree $d\geq 2$ and let $\widetilde{X}$ be a compact subset of the space of rational maps of $\widehat{\mathbf{C}}$ of degree d such that $Q_i\in\widetilde{X}$ for all i. Let $X=\widetilde{X}^{\mathbf{N}}$ and define $g: X\to X$ by the shift $g((R_i))=(R_{i+1})$. Then the map $f: X\times\widehat{\mathbf{C}}\circlearrowleft$ defined by $f((R_i),z)=((R_{i+1}),R_0(z))$ is a fibered rational map over g, and the restriction of f to $\{(Q_i)\}\times\widehat{\mathbf{C}}$ can be identified with the sequence (Q_i). For more on random and non-autonomous dynamical systems see e.g. [K], [KS] and [Bo]. The papers [FS1] and [FW] are concerned with random iterations on $\widehat{\mathbf{C}}$ and $\mathbf{P}^2$, respectively.

Some of the results in this paper generalize to the setting of fibered holomorphic mappings of $\mathbf{P}^k$, $k\geq 1$, that is, replacing $\widehat{\mathbf{C}}\simeq\mathbf{P}^1$ by $\mathbf{P}^k$. We will not, however, pursue this generalization here. At any rate, the proof of the uniqueness result in Theorem B(iii) does not generalize. Indeed, as mentioned above it is an open problem whether uniqueness holds even in the non-fibered case for $k\geq 2$.

The organization of this paper is as follows. The definition of $\widehat{\mathbf{C}}$-bundles and fibered rational maps are given in the first section. In Section 2 we show how to define natural measures μ_x on the fibers $Y_x=\pi^{-1}(x)$, using pluripotential theory. The support of μ_x is equal to the Julia set for the restriction of $\{f^n\}$ to Y_x. In the case when the base space X is a single point we recover the unique measure of maximal entropy in Theorem A. We compute topological entropy of fibered rational maps in Section 3. Given an invariant measure μ' for $g: X\circlearrowleft$ we can construct a measure μ on Y having μ_x as conditional measures on the fibers Y_x and such that $\pi_*\mu=\mu'$. The properties of μ are studied in Sections 4–6, where we show that μ is ergodic (mixing) if μ' is. Further, μ is the unique measure of maximal entropy among the invariant measures ν for f such that $\pi_*\nu=\mu'$. Finally, in Section 7 we apply the previous results to selfmaps of ruled surfaces and prove Theorems C and D.

1. Fibered rational maps

The phase space for the dynamical systems in this paper will be a space Y, fibered over another space X with Riemann spheres as fibers. The dynamical systems themselves will be continuous selfmaps of Y mapping fibers to fibers as rational mappings. The purpose of this section is to define all of this in a precise way.

Throughout the paper, $\widehat{\mathbf{C}}$ will denote the Riemann sphere, equipped with the spherical metric. The metric is normalized so that the diameter of $\widehat{\mathbf{C}}$ is one.

Definition 1.1. Let X be a compact metric space, and let $\pi: Y\to X$ be a $\widehat{\mathbf{C}}$-bundle. This means that $Y_x:=\pi^{-1}(x)$ is homeomorphic to a sphere for every x, and that Y_x comes with a complex structure and a smooth $(1,1)$-form $\omega_x>0$ inducing the metric on Y_x. We also assume that X can be covered by open sets $\{U_i\}$ for which there exists a homeomorphism $\Phi_i: U_i\times\widehat{\mathbf{C}}\to\pi^{-1}(U_i)$ and such that $\Phi_j^{-1}\circ\Phi_i: (U_i\cap U_j)\times\widehat{\mathbf{C}}\circlearrowleft$ maps $\{x\}\times\widehat{\mathbf{C}}$ as a Möbius transformation onto itself for all $x\in U_i\cap U_j$.

Remark 1.2. The definition implies that given $x_0\in X$ we may find a continuous family $i_x: \widehat{\mathbf{C}}\to Y_x$ of conformal mappings for x close to x_0. Such a family i_x will be called a *local parameterization.* The local parameterizations are not uniquely defined, but by compactness of X we may assume that there exists a compact subset M_0 of the set of Möbius transformations of $\widehat{\mathbf{C}}$ such $i_x\circ j_x^{-1}\in M_0$ for any two local parameterizations i_x and j_x. As we will see, the choice of local parameterization will not be important for most of what follows. If i_x is a local parameterization, then the pullback $i_x^*\omega_x$ is a positive smooth form on $\widehat{\mathbf{C}}$, depending continuously on x.

Example 1.3. Let $Y=X\times\widehat{\mathbf{C}}$ and let π be the projection on the first coordinate.

Example 1.4. Let Y be a ruled surface over a Riemann surface X. This means that Y is a smooth projective variety of complex dimension 2, which is also a holomorphic $\mathbf{P}^1$-bundle over X. Every Y_x has then a unique conformal structure and a positive form $\omega_x = \omega|_{Y_x}$, where ω is the Kähler form on Y.

The following result will be needed in Section 4. It says that continuous functions on $\widehat{\mathbf{C}}$-bundles can be approximated by functions which are smooth on the fibers. The proof, using a partition of unity on X, is left to the reader.

Lemma 1.5. *Let Y be a $\widehat{\mathbf{C}}$-bundle over X and let $\varphi \in C^0(Y)$. Then, given $\varepsilon > 0$ there exists $C > 0$ and $\widetilde{\varphi} \in C^0(Y)$ such that $\|\widetilde{\varphi} - \varphi\|_\infty < \varepsilon$, $\widetilde{\varphi}_x := \widetilde{\varphi}|_{Y_x} \in C^2(Y_x)$ and $\|\widetilde{\varphi}_x\|_{C^2(Y_x)} \leq C$ for $x \in X$.*

We next define fibered rational maps.

Definition 1.6. Let Y be a $\widehat{\mathbf{C}}$-bundle over X and let $g: X \circlearrowleft$ be continuous. We say that a continuous mapping $f: Y \circlearrowleft$ is a *rational map of degree d, fibered over g* if it has the following properties:

(i) π semiconjugates f to g, i.e. $g \circ \pi = \pi \circ f$, in other words, f maps the fiber Y_x into the the fiber $Y_{g(x)}$ for $x \in X$;

(ii) the restriction $f|_{Y_x}: Y_x \to Y_{g(x)}$ is a rational map of degree d.

Condition (ii) can also be phrased as follows: for any local parameterizations i_x at x and $i_{g(x)}$ at $g(x)$, the composition $i_{g(x)}^{-1} \circ f \circ i_x$ is a rational map of $\widehat{\mathbf{C}}$ of degree d.

If $f: Y \circlearrowleft$ is a rational map, fibered over $g: X \circlearrowleft$, and $x \in X$, $y \in Y$, $n \geq 0$, then we will write x_n for $g^n(x)$ and y_n for $f^n(y)$. We will denote the restriction of f to Y_x by f_x. Similarly, f_x^n is the restriction of f^n to Y_x.

Example 1.7. Let $Y = X \times \widehat{\mathbf{C}}$ as in Example 1.3. A rational map $f: Y \circlearrowleft$ of degree d, fibered over $g: X \circlearrowleft$ is then of the form

$$f(x, z) = (g(x), Q_x(z)),$$

where Q_x is a rational function of degree d, depending continuously on x. A special case is when the mappings Q_x are polynomial mappings of $\mathbf{C}$. Such mappings have been studied by Sester [S1], [S2]; see also [H2] and [J].

Example 1.8. Let Y be a ruled surface over X as in Example 1.4. It is a result of Dabija [D] that (almost) every holomorphic selfmap of Y is a rational map fibered over a holomorphic map $g: X \circlearrowleft$. See Proposition 7.1 for more details.

2. Invariant measures and Julia sets on fibers

In this section we construct probability measures on the fibers of a fibered rational map. The support of these measures serve as Julia sets for the restriction of the dynamics to the fibers.

Let $f: Y\circlearrowleft$ be a rational map of degree $d\geq 2$, fibered over $g: X\circlearrowleft$. The form ω_x on Y_x induces a measure, also called ω_x on Y_x, or even on Y. As measures on Y we have that $x\mapsto\omega_x$ is weakly continuous.

If χ is a continuous function on Y_x, then we define the continuous function $(f_x^n)_*\chi$ on Y_{x_n} by

$$(f_x^n)_*\chi(z)=\sum_{f_x^n(w)=z}\chi(w),$$

where the preimages w of z are counted with multiplicity. We define pullbacks of measures by duality, i.e. $\langle (f_x^n)^*\nu,\chi\rangle=\langle\nu,(f_x^n)_*\chi\rangle$. Now define probability measures $\mu_{x,n}$ on Y_x by

$$\mu_{x,n}:=\frac{1}{d^n}(f_x^n)^*\omega_{x_n}.$$

Theorem 2.1. *The measures $\mu_{x,n}$ converge weakly to a probability measure μ_x on Y_x. Further:*

(i) *μ_x puts no mass on polar subsets of Y_x;*

(ii) *$(f_x)_*\mu_x=\mu_{x_1}$ and $(f_x)^*\mu_{x_1}=d\cdot\mu_x$;*

(iii) *$x\mapsto\mu_x$ is continuous in the weak topology of measures on Y.*

We will prove Theorem 2.1 by finding potentials for the measures μ_x and proving convergence results for these potentials. This method was first used by Hubbard and Papadopol [HP] and further developed by Fornæss and Sibony (see [FS3]). The main idea is to lift $f_x: Y_x\to Y_{x_1}$ to a selfmap of $\mathbf{C}^2$; the fact that there is, in general, no canonical way of doing this makes the proof below slightly technical.

Proof. Throughout the proof, $C>0$ will denote constants not depending on x, n, or any choice of local parameterization. Let $\mathbf{C}^2_*=\mathbf{C}^2\setminus\{0\}$ and let $\pi':\mathbf{C}^2_*\to\widehat{\mathbf{C}}\simeq\mathbf{P}^1$ be the natural projection. Any Borel probability measure ν on $\widehat{\mathbf{C}}$ can be identified with a plurisubharmonic function G_ν on $\mathbf{C}^2_*$ such that $G_\nu(z,w)\leq\log|(z,w)|+O(1)$, as $|(z,w)|\to\infty$, and $G_\nu(\lambda z,\lambda w)=G_\nu(z,w)+\log|\lambda|$ for $\lambda\in\mathbf{C}^*$. The function G_ν is unique up to an additive constant and the correspondence is given by $\nu=dd^c(G_\nu\circ s)$, where s is any local section of π' and $d^c=(\sqrt{-1}/2\pi)(\bar\partial-\partial)$.

In our setting, given a local parameterization $i_x:\widehat{\mathbf{C}}\to Y_x$ there exists a smooth potential $G_{x,0}$ for ω_x in the sense that $\omega_x=dd^c(G_{x,0}\circ s\circ i_x^{-1})$. By the smoothness of ω_x we may assume that

$$\log|(z,w)|\leq G_{x,0}(z,w)\leq\log|(z,w)|+C\quad\text{for } (z,w)\in\mathbf{C}^2_*. \tag{2.1}$$

Thus, if $G_{x,0}$ and $\widetilde{G}_{x,0}$ are two different potentials, then we may assume that

$$|\widetilde{G}_{x,0}(z,w)-G_{x,0}(z,w)|\leq C. \tag{2.2}$$

Given a local parameterization at x we may assume that $x \mapsto G_{x,0}$ is uniformly continuous.

Given $x \in X$ we will lift $f_x: Y_x \to Y_{x_1}$ to selfmaps of $\widehat{\mathbf{C}}$ and $\mathbf{C}^2_*$. Let i_x and i_{x_1} be local parameterizations near x and x_1, Define $Q_x: \widehat{\mathbf{C}} \circlearrowleft$ to be a rational map and $R_x: \mathbf{C}^2_* \circlearrowleft$ to be a homogeneous polynomial map, both of degree d, such that

$$\sup\{|R_x(z,w)| : |(z,w)| = 1\} = 1$$

and such that the diagram

$$\begin{array}{ccccc} \mathbf{C}^2_* & \xrightarrow{\pi'} & \widehat{\mathbf{C}} & \xrightarrow{i_x} & Y_x \\ \Big\downarrow{\scriptstyle R_x} & & \Big\downarrow{\scriptstyle Q_x} & & \Big\downarrow{\scriptstyle f_x} \\ \mathbf{C}^2_* & \xrightarrow{\pi'} & \widehat{\mathbf{C}} & \xrightarrow{i_{x_1}} & Y_{x_1} \end{array} \tag{2.3}$$

commutes. Given the local parameterizations i_x and i_{x_1} these properties determine Q_x uniquely, and R_x uniquely up to multiplication by a complex number of unit modulus. Further, if we change i_{x_1} and $\widetilde{R}_x$ is the resulting map of $\mathbf{C}^2_*$, then we have

$$e^{-C} \leq \frac{|\widetilde{R}_x(z,w)|}{|R_x(z,w)|} \leq e^C. \tag{2.4}$$

Now consider an orbit $(x_j)_{j \geq 0}$ in X, select parameterizations at each point x_j, and let R_{x_j} be the corresponding homogeneous selfmaps of $\mathbf{C}^2_*$. Let R^n_x be the composition $R_{x_{n-1}} \circ \ldots \circ R_x$. Then R^n_x is a homogeneous polynomial mapping of $\mathbf{C}^2_*$ of degree d^n. Notice that R^n_x is determined, up to multiplication by a complex number of unit modulus, by the local parameterizations at x and x_n. Given these choices we have that $x \mapsto R^n_x$ is continuous. If we change i_{x_n}, then, corresponding to (2.4) we have

$$e^{-C} \leq \frac{|\widetilde{R}^n_x(z,w)|}{|R^n_x(z,w)|} \leq e^C. \tag{2.5}$$

Define the plurisubharmonic function $G_{x,n}$ on $\mathbf{C}^2_*$ by

$$G_{x,n} = \frac{1}{d^n} G_{x_n,0} \circ R^n_x.$$

Then $G_{x,n}$ is uniquely defined, given the local parameterizations at x_0 and x_n. Further, $x \mapsto G_{x,n}$ is uniformly continuous near x_0. From (2.1), (2.2) and (2.5) it follows that if we change the local parameterizations at x_n, and $\widetilde{G}_{x,n}$ is the modified $G_{x,n}$, then

$$|G_{x,n} - \widetilde{G}_{x,n}| \leq \frac{C}{d^n}. \tag{2.6}$$

Lemma 2.2. *Given $x \in X$ and a local parameterization at x, the functions $G_{x,n}$ converge uniformly on compact subsets of $\mathbf{C}^2_*$ to a plurisubharmonic function G_x, as $n \to \infty$. This function does not depend on the choice of local parameterizations at x_j for $j \geq 1$, and it satisfies $G_x(z,w) \leq \log|(z,w)| + O(1)$, as $|(z,w)| \to \infty$, and $G_x(\lambda z, \lambda w) = G_x(z,w) + \log|\lambda|$ for $\lambda \in \mathbf{C}^*$. Further, we have:*

(i) *G_x is uniformly continuous on $\mathbf{C}^2_*$;*

(ii) *$G_{x_1} \circ R_x = d \cdot G_x$, given i_x and i_{x_1};*

(iii) *$x \mapsto G_x$ is uniformly continuous.*

Proof. The fact that the choice of local parameterizations at x_j, $j \geq 1$, is irrelevant follows from (2.6). Let us therefore fix local parameterizations at these points. Notice that

$$e^{-C}|(z,w)|^d \leq |R_{x_j}(z,w)| \leq |(z,w)|^d$$

for $j \geq 0$ and $(z,w) \in \mathbf{C}^2_*$. This implies that

$$e^{-C}|R^n_x(z,w)|^d \leq |R^{n+1}_x(z,w)| \leq |R^n_x(z,w)|^d,$$

so, using (2.1),

$$|G_{x,n+1} - G_{x,n}| \leq \frac{C}{d^n},$$

which shows that $G_x := \lim_{n \to \infty} G_{x,n}$ exists and that

$$|G_{x,n} - G_x| \leq \frac{C}{d^n}. \tag{2.7}$$

That G_x is continuous and plurisubharmonic, that $G_x(\lambda z, \lambda w) = G_x(z,w) + \log|\lambda|$ and that $G_x(z,w) \leq |(z,w)| + O(1)$ all follow from the corresponding properties of $G_{x,n}$ and from (2.7).

We see from the definition of $G_{x,n}$ that $G_{x_1,n} \circ R_x = d \cdot G_{x,n+1}$. This and (2.7) imply (ii). Finally (iii) follows from the estimate (2.7) and the fact that $x \mapsto G_{x,n}$ is continuous for every n. $\square$

We now continue the proof of Theorem 2.1. First we note that $G_{x,n}$ is a potential for $\mu_{x,n}$ in the sense that $\mu_{x,n} = dd^c(G_{x,n} \circ s \circ i_x^{-1})$ for a local section s of π'. Indeed, for any local section t of π' we have that the function $d^{-n} G_{x_n,0} \circ t \circ i_{x_n}^{-1} \circ f^n_x$ is a potential for $\mu_{x,n}$. Further, $t \circ i_{x_n}^{-1} \circ f^n_x$ differs from $R^n_x \circ s \circ i_x^{-1}$ by a holomorphic factor $\phi \neq 0$. Thus

$$dd^c(d^{-n} G_{x,0} \circ R^n_x \circ s \circ i_x^{-1}) = dd^c(d^{-n} \log|\phi|) + dd^c(d^{-n} G_{x_n,0} \circ t \circ i_{x_n}^{-1} \circ f^n_x) = 0 + \mu_{x,n}.$$

Since $G_{x,n} \to G_x$ uniformly on compact subsets of $\mathbf{C}^2_*$, as $n \to \infty$, it follows that $\mu_{x,n}$ converges weakly to a probability measure μ_x on Y_x. Here $\mu_x = dd^c(G_x \circ s \circ i_x^{-1})$ for a local section s of π'. The properties (i)–(iii) of μ_x now follow from the corresponding properties of G_x. $\square$

We next show that the support of the measure μ_x can be interpreted as the Julia set of f in Y_x. More precisely, fix $x \in X$ and consider the family $\{f_x^n\}_{n\geq 0}$ of rational maps on Y_x. We define F_x, the *Fatou set of f in Y_x*, to be the set where this family is normal. Equivalently, $z\in F_x$ if and only if the family $Q_x^n = i_{x_n}^{-1} \circ f_x^n \circ i_x$ of rational maps on $\widehat{\mathbf{C}}$ is normal near $i_x^{-1}(z)$; this does not depend on the choices local parameterizations at x and x_n. Still equivalently, F_x is the open subset of Y_x where the family $\{f_x^n\}$ of mappings from Y_x into Y is locally equicontinuous. Clearly F_x is an open subset of Y_x. Its complement $J_x := Y_x \setminus F_x$ is called the *Julia set of f in Y_x*.

Proposition 2.3. *The support of μ_x is equal to J_x for every $x\in X$.*

Proof. Again we let $C>0$ denote a constant not depending on x, n, or any choice of local parameterization. We will follow the proof of Theorem 6.4 in [FS3]. Fix $x\in X$ and local parameterizations i_x at x and i_{x_n} at x_n for $n\geq 1$. Define $Q_x^n = Q_{x_{n-1}} \circ ... \circ Q_x$ and $R_x^n = R_{x_{n-1}} \circ ... \circ R_x$ using (2.3). Let us first show that the support of μ_x is contained in J_x. Suppose that $U \Subset i_x^{-1}(F_x)$ and let $Q_x^{n_j}$ be a subsequence converging uniformly to a meromorphic function on U. After shrinking U, if necessary, we may assume without loss of generality that $Q_x^{n_j}(U) \subset \{[z:w] : |z| \leq |w|\}$ for all j. Then $R_x^{n_j}(z,w) = \varrho_j(z,w)(A_j(z,w),1)$, where ϱ_j and A_j are holomorphic on $(\pi')^{-1}(U)$ and $\varrho_j \neq 0$, $|A_j|\leq 1$. Thus

$$\frac{1}{d^{n_j}} \log |R_x^{n_j}(z,w)| = \frac{1}{d^{n_j}} \log |\varrho_j(z,w)| + \frac{1}{d^{n_j}} \log |(A_j(z,w),1)|.$$

Here the left-hand side converges uniformly to G_x by Lemma 2.2, (2.1) and (2.2). Further, the first term in the right-hand side is pluriharmonic, and the last term converges uniformly to 0. Thus G_x is pluriharmonic on $(\pi')^{-1}(U)$, so $i_x(U) \cap \operatorname{supp} \mu_x = \emptyset$. It follows that $\operatorname{supp} \mu_x \subset J_x$.

For the converse, suppose that μ_x vanishes on an open set $i_x(U)$ so that G_x is pluriharmonic on $(\pi')^{-1}(U)$. By shrinking U we may assume that there exists a holomorphic function $h\neq 0$ on $(\pi')^{-1}(U)$ such that $G_x = \log |h|$. But then it follows from (2.7) and (2.1) that

$$\left| \frac{1}{d^n} \log |R_x^n| - \log |h| \right| \leq \frac{C}{d^n}$$

on $(\pi')^{-1}(U)$. Thus

$$e^{-C} \leq \left| \frac{R_x^n}{h^{d^n}} \right| \leq e^C$$

on $(\pi')^{-1}(U)$, which implies that R_x^n/h^{d^n} is normal on $(\pi')^{-1}(U)$. Since π' semiconjugates R_x^n/h^{d^n} to Q_x^n, it follows that $\{Q_x^n\}$ is normal on U. $\square$

The following proposition sums up some of the properties of the Julia sets J_x. They follow from the corresponding properties of μ_x.

Proposition 2.4. *Let $f: Y\circlearrowleft$ be a rational map of degree $d\geq 2$ fibered over $g: X\circlearrowleft$. Then the Julia set J_x of f in Y_x has the following properties:*

(i) *J_x is compact, nonempty and has no isolated points for $x\in X$;*

(ii) *$f_x J_x = J_{x_1}$ and $f_x^{-1} J_{x_1} = J_x$ for $x\in X$;*

(iii) *the assignment $x\mapsto J_x$ is lower semicontinuous in the Hausdorff metric on compact subsets of Y.*

Proof. (i) This follows because μ_x is a probability measure and μ_x has continuous local potentials, given by $G_x\circ s\circ i_x^{-1}$, where s is a local section of $\pi': \mathbf{C}_*^2\to\widehat{\mathbf{C}}$.

(ii) This follows from $(f_x)_*\mu_x = \mu_{x_1}$ and $f_x^*\mu_{x_1} = d\cdot\mu_x$.

(iii) This is a consequence of the continuity of $x\mapsto\mu_x$. $\square$

Remark 2.5. The assignment $x\mapsto J_x$ is not continuous, in general. See e.g. [J] for examples. This is analogous to the fact that the Julia set of a rational function depending on a parameter generally does not vary continuously with the parameter.

The measure μ_x was defined by pulling back the measure ω_{x_n} on Y_{x_n} by f_x^n, normalizing, and letting $n\to\infty$. One may ask what happens if we pull back other measures. The following result asserts that for most points $w\in Y_{x_n}$, the preimages of w under f_x^n are distributed like μ_x. The proof is almost identical to the proof of Lemma 8.3 in [FS3]; the changes needed are left to the industrious reader. We will use Proposition 2.6 in Section 4.

Proposition 2.6. *There exists a constant $C>0$ such that if $x\in X$ and φ is a continuous function on Y_x such that $\varphi_x := \varphi|_{Y_x}\in C^2(Y_x)$, then*

$$\mu_{x_n}\left\{w\in Y_{x_n} : \left|\frac{1}{d^n}(f_x^n)_*\varphi(w) - \langle\mu_x,\varphi\rangle\right| > t\right\} \leq \frac{C|\varphi_x|_{C^2(Y_x)}}{td^n}$$

for $t>0$ and $n\geq 1$.

3. Topological entropy

In the next four sections we will study the ergodic properties of fibered rational maps. For background on ergodic theory see [W]. First we will consider topological entropy. The exposition follows Lyubich [L].

We recall the definition of topological entropy due to Bowen [Bo]. Let (Y, d) be a compact metric space and $f\colon Y\circlearrowleft$ a continuous mapping. A set $Z\subset Y$ is (n,δ)-*separated* if for every two distinct points $z, w\in Z$ there exists i such that $0\le i<n$ and $d(f^iz, f^iw)>\delta$. A $(1,\delta)$-separated set will also be called a δ-*net* (this notion does not require a map f). A set $F\subset Y$ (n,δ)-*spans* another set $Z\subset Y$ if for every $z\in Z$ there exists $w\in F$ such that $d(f^iz, f^iw)<\delta$ for $0\le i<n$.

For a compact set $Z\subset Y$ let $r_n(\delta, Z)$ be the smallest cardinality of any set F which (n,δ)-spans Z, and let $s_n(\delta, Z)$ be the largest cardinality of any (n,δ)-separated subset of Z. It is then easy to see that

$$r_n(\delta, Z)\le s_n(\delta, Z)\le r_n\left(\tfrac{1}{2}\delta, Z\right)<\infty.$$

Also, $r_n(\delta, Z)$ and $s_n(\delta, Z)$ are decreasing in δ. Thus it makes sense to define

$$h(f, Z)=\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{n}\log r_n(\delta, Z)=\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{n}\log s_n(\delta, Z).$$

The number $h(f, Z)$ is called the topological entropy of f on Z. If we want to emphasize the dependence on f, then we will write e.g. $s_n(\delta, Z; f)$. On the other hand, if $Z=Y$, then we will suppress Y and write $s_n(\delta; f)$ and $h(f)$.

The following theorem is the main result in this section.

Theorem 3.1. *If $f\colon Y\circlearrowleft$ is a rational map of degree d, fibered over $g\colon X\circlearrowleft$, then the following holds*:

(i) $h(f, Y_x)=\log d$ *for every* $x\in X$;

(ii) $h(f)=h(g)+\log d$.

Note that if X is a single point, then we recover the result by Gromov [G] and Lyubich [L] that the topological entropy of a rational map of $\widehat{\mathbf{C}}$ of degree d is $\log d$.

Theorem 3.1 follows from the following two results.

Lemma 3.2. *If $f\colon Y\circlearrowleft$ is a rational map of degree d, fibered over $g\colon X\circlearrowleft$, then for every $\alpha\in(0,1)$ there exists $\delta_0=\delta_0(\alpha)>0$ with the property that for $\delta\le\delta_0$, $n\ge 1$ and $x\in X$ there exists an (n,δ)-separating set in Y_x with at least $d^{\alpha n}$ elements.*

Lemma 3.3. *If $f\colon Y\circlearrowleft$ is a rational map of degree d, fibered over $g\colon X\circlearrowleft$, then for every $\delta>0$ there exists a constant $C(\delta)>0$ with the property that for every $x\in X$ and every $n\ge 1$ there exists an (n,δ)-spanning set in Y_x with at most $C(\delta)n^5d^n$ elements.*

We postpone the proofs of the above lemmas for a moment, and show how they imply Theorem 3.1.

Proof of Theorem 3.1. (i) Fix $x \in X$. From Lemma 3.2 it follows that

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{n}\log s_n(\delta, Y_x)\geq \alpha\log d$$

for every $\alpha<1$, so $h(f,Y_x)\geq\log d$. On the other hand, Lemma 3.3 implies that

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{n}\log r_n(\delta, Y_x)\leq \lim_{\delta\to 0}\limsup_{n\to\infty}\frac{1}{n}\log C(\delta)n^5d^n=\log d.$$

Thus $h(f,Y_x)\leq\log d$, so $h(f,Y_x)=\log d$.

(ii) By a result of Bowen [Bo, Theorem 17] we have

$$h(f)\leq h(g)+\sup_{x\in X}h(f,Y_x),$$

so by (i) we get $h(f)\leq h(g)+\log d$. On the other hand, fix $\alpha\in(0,1)$, $\delta\leq\delta_0$, with δ_0 from Lemma 3.2, and $n\geq 1$. Find $\delta'>0$ such that $d_Y(Y_x,Y_{x'})>\delta$ whenever $d_X(x,x')>\delta'$. Let $E\subset X$ be a maximal (n,δ')-separating set with respect to g and, for $x\in E$, let F_x be a maximal (n,δ)-separating subset of Y_x with respect to f. By Lemma 3.2, F_x has at least $d^{\alpha n}$ elements for every $x\in E$. Let $F=\bigcup_{x\in E}F_x$. Then F is (n,δ)-separating for f so

$$s_n(\delta;f)\geq|F|\geq s_n(\delta';g)d^{\alpha n}.$$

It follows easily from this that $h(f)\geq h(g)+\log d$. Thus $h(f)=h(g)+\log d$ and we are done. □

We now give the proofs of Lemmas 3.2 and 3.3. The proof of Lemma 3.2 is an adaptation of an elegant argument by Misiurewicz and Przytycki [MP].

Proof of Lemma 3.2. Let $L=\max(2,\sup_{x\in X}|Df_x|)$ and let $\varepsilon=L^{-\alpha/(1-\alpha)}$. Fix $x\in X$ and write $Y_i=Y_{x_i}$, $f_i=f_{x_i}$ and $f_i^k=f_{x_i}^{k-i}$ for $0\leq i\leq k$.

Define $B_i:=\{z\in Y_i:|Df_i(z)|\geq\varepsilon\}$. Then B_i is compact and since the second derivative of f_i is uniformly bounded, there exists $\delta_0>0$, not depending on x or i, such that f_i is univalent on every disk of radius δ_0 centered at points in B_i. This implies that if $z_1,z_2\in B_i$ and $d(z_1,z_2)<\delta_0$, then $f_i(z_1)\neq f_i(z_2)$ unless $z_1\neq z_2$.

Fix $n\geq 1$, $\delta\leq\delta_0$ and let

$$A=\{z\in Y_0:|\{i:0\leq i<n,\ f_0^i(z)\in B_i\}|\leq\alpha n\}.$$

Then $Y_n\setminus f_0^n(A)$ has positive Lebesgue measure. Indeed, if $z\in A$, then

$$|Df_0^n(z)|<L^{\alpha n}\varepsilon^{(1-\alpha)n}=1.$$

Thus we may find a point $z\in Y_n$ which is not in $f_0^n(A)$ and is also not a critical value of f_0^n. We will define our (n,δ)-separating set as a suitable subset of $(f_0^n)^{-1}(z)$.

First, if $w\in Y_{i+1}$ and w is not a critical value of f_i, then define $R_i(w)\subset Y_i$ as follows: if $f_i^{-1}(w)\subset B_i$, then $R_i(w)=f_i^{-1}(w)$; otherwise $R_i(w)$ is any point in $f_i^{-1}(w)\setminus B_i$.

Now define the sets $S_i\subset Y_i$, $0\le i\le n$ inductively by

$$S_n=\{z\},\quad S_i=\bigcup_{w\in S_{i+1}} R_i(w) \text{ for } 0\le i<n.$$

It is easy to check, inductively, that S_i is $(n-i,\delta)$-separated. In particular S_0 is (n,δ)-separated.

We claim that S_0 has at least $d^{\alpha n}$ elements. To see this, let $m=[\alpha n]+1$ and consider the set T of pairs (w,i) such that $0\le i\le n$, $w\in S_i$ and $f_i^{j+i}(w)\in S_{j+i}$ for exactly m numbers j, $0\le j\le n-i$. A combinatorial argument shows that T has exactly d^m elements. But there are at least as many points in S_0 as there are elements in T. Thus $|S_0|\ge d^m\ge d^{\alpha n}$.

Hence we have found an (n,δ)-separating subset S_0 of Y_x with at least $d^{\alpha n}$ elements. This completes the proof. □

Finally we turn to the proof of Lemma 3.3. In the case when X is a single point, the result is due to Lyubich [L]. The proof below is very similar to Lyubich's.

We will make use of two results which have nothing to do with dynamics. They are consequences of the Koebe distortion theorem and the geometry of $\widehat{\mathbf{C}}$, respectively. The first result is the following

Lemma 3.4. (Proposition 8 in [L]) *Given $0<\eta<\frac{1}{2}$ and $\delta>0$ there exists $\varkappa=\varkappa(\eta,\delta)\in(0,1)$ with the property that if*

$$h\colon B(u,\varrho)\longrightarrow\widehat{\mathbf{C}},\quad 0<\varrho<1,$$

is a univalent meromorphic function which avoids some η-net on $\widehat{\mathbf{C}}$, then

$$hB(u,\varkappa\varrho)\subset B(hu,\delta).$$

Proof of Lemma 3.3. Now fix $\eta\in\left(0,\frac{1}{2}\right)$ and $\delta>0$ and find $\varkappa=\varkappa(\eta,\delta)\in(0,1)$ as in Lemma 3.4. Fix $x\in X$, keeping in mind that all the estimates below will be uniform in x. For $0\le k\le i$, write $Y_i=Y_{x_i}$, $f_i=f_{x_i}$ and $f_k^i=f_{x_k}^{i-k}$. Let Z_i' be a finite η-net in Y_i containing all the critical points of f_i. We assume that $m_0:=\sup_i|Z_i'|<\infty$. For $i>0$ let Z_i be the subset of Y_i defined by

$$Z_i=\bigcup_{0\le k\le i} f_k^i(Z_k').$$

Then $|Z_i| \leq (i+1)m_0$ for all i. Notice that if $w \in Y_i \setminus Z_i$ and $0 \leq k < i$, then all branches of $(f_k^i)^{-1}$ are single-valued near w and do not take any value in Z'_k.

Fix $n \geq 1$ and let $\varepsilon = \delta/2L^n$, where $L = \sup_{i,z} |Df_i(z)|$. We now invoke the second result referred to above. Namely, by a geometrical argument [L, Proposition 7], there exist finite subsets $A_i \subset Y_i \setminus Z_i$ for $0 \leq i \leq n$ such that

(i) $|A_i| \leq C(\delta) n^2 (\log(\delta/\varepsilon) + C(\delta)) \leq C(\delta) n^3$;

(ii) for every $z \in Y_i$ with $d(z, Z_i) \geq \varepsilon$ there exists $u \in A_i$ such that $d(u, z) < \varrho(u)$, where $\varrho(u) = \min\big(\varkappa d(u, Z_i), \tfrac{1}{2}\delta\big)$;

(iii) A_i contains no critical values of f_k^i for $0 \leq k < i$, further, if Ω_i is a given open dense subset of Y_i, then $A_i \cap \Omega_i = \emptyset$.

Let $B_i = (f_x^i)^{-1}(A_i)$. Then $B_i \subset Y_x$ and $|B_i| = d^i |A_i|$. Fix $z \in Y_x$ and consider the orbit $(z_i)_{0 \leq i \leq n}$ of length n, where $z_i = f_x^i z$. We consider three cases.

The first case is when there exists m, $0 \leq m < n$, such that $d(z_i, Z_i) \geq \varepsilon$ for $0 \leq i \leq m$ but $d(z_{m+1}, Z_{m+1}) < \varepsilon$. Then pick a point $u \in A_m$ such that $d(z_m, u) < \varrho(u)$ and a point $w \in Z_{m+1}$ such that $d(z_{m+1}, w) < \varepsilon$. Since $\varrho(u) \leq \varkappa d(u, Z_m)$, all branches of $(f_i^m)^{-1}$ are single-valued on the disk $B(u, \varrho(u)/\varkappa)$ for $0 \leq i < m$. Let $g_{i,m}$ be the branch mapping z_m to z_i and let $v = g_{0,m}(u) \in B_m$. The map $g_{i,m}$ avoids the η-net Z'_i, so we may apply Lemma 3.4 to $g_{i,m}$ and conclude that

$$d(z_i, f_x^i v) = d(g_{i,m}(z_m), g_{i,m}(u)) < \tfrac{1}{2}\delta \quad \text{for } 0 \leq i < m. \tag{3.1}$$

Moreover, since $\varrho(u) < \frac{1}{2}\delta$ we have

$$d(z_m, f_x^m v) = d(z_m, u) < \tfrac{1}{2}\delta. \tag{3.2}$$

On the other hand, by the choice of ε and L we have

$$d(z_i, f_{m+1}^i w) \leq L^{i-m-1}\varepsilon < \tfrac{1}{2}\delta \quad \text{for } m < i \leq n. \tag{3.3}$$

The second case is when $d(z_i, Z_i) \geq \varepsilon$ for $0 \leq i \leq n$ $(m = n)$. Then there exists a point $v \in B_n$ such that

$$d(z_i, f_x^i v) < \tfrac{1}{2}\delta \quad \text{for } 0 \leq i \leq n.$$

The third case is when $d(z, Z_0) < \varepsilon$ $(m = -1)$. Then there exists $w \in Z_0$ such that

$$d(z_i, f_x^i w) < \tfrac{1}{2}\delta \quad \text{for } 0 \leq i \leq n.$$

For each triple (m, v, w) with $-1 \leq m \leq n$, $v \in B_m$ and $w \in Z_{m+1}$ pick a point $z(m, v, w)$ such that (3.1)–(3.3) hold, if such a point exists. Let F_x be the set of all the chosen points $z(m, v, w)$. It then follows that F_x (n, δ)-spans Y_x. Now there are at most $n+1$ choices for m, $d^n C(\delta) n^3$ choices for v and $(n+1)|Z'| \leq Cn$ choices for w. Thus F_x has at most $C(\delta) n^5 d^n$ elements. This completes the proof. $\square$

4. Mixing properties of μ

Let $f: Y \circlearrowleft$ be a rational map of degree $d \geq 2$, fibered over $g: X \circlearrowleft$. In Section 2 we constructed measures μ_x on the fibers Y_x with certain invariance properties. Given a probability measure μ' for g we can define a probability measure μ on Y by

$$(4.1) \qquad \langle \mu, \varphi \rangle := \int_X \left(\int_{Y_x} \varphi \, \mu_x \right) \mu'(x)$$

for continuous functions φ on Y. Thus $\pi_* \mu = \mu'$ and the conditional measures of μ on the fibers Y_x are given by μ_x. Conversely, these two properties define μ uniquely.

In the following three sections we will study the dynamical properties of μ. Here we investigate when μ is ergodic or mixing. We start by the following simple result.

Proposition 4.1. *If μ' is invariant for g, then μ is invariant for f.*

Proof. We have

$$\langle f_* \mu, \varphi \rangle = \langle \mu, \varphi \circ f \rangle = \int_X \langle \mu_x, \varphi \circ f_x \rangle \, \mu'(x) = \int_X \langle \mu_{x_1}, \varphi \rangle \, \mu'(x) = \langle \mu, \varphi \rangle.$$

The third equality follows from $(f_x)_* \mu_x = \mu_{x_1}$ and the fourth equality from the invariance of μ'. $\square$

We now turn to the ergodic properties of μ. In the case when X is a single point, then $Y \simeq X$ and f is essentially a rational map of $\widehat{\mathbf{C}}$. It is a result of Lyubich [L] and of Freire, Lopez and Mañé [FLM] that the system (f, μ) is exact in this case. In particular, μ is ergodic and (strongly) mixing for f. The following result generalizes this to general fibered rational maps.

Theorem 4.2. *Let $f: Y \circlearrowleft$ be a rational map of degree $d \geq 2$, fibered over $g: X \circlearrowleft$, and let μ' be an invariant probability measure for g. Define the invariant measure μ for f by* (4.1). *Then the following holds*:

(i) *if μ' is ergodic, then so is μ;*

(ii) *if μ' is (strongly) mixing, then so is μ.*

Proof. We will prove (ii) and then indicate how to handle (i). The proof below follows the proof of Theorem 8.2 in [FS3]. Thus suppose that μ' is mixing for $g: X \circlearrowleft$. To show that μ is mixing for f, it is sufficient to show that

$$(4.2) \qquad \langle \mu, \varphi(\psi \circ f^n) \rangle \to \langle \mu, \varphi \rangle \langle \mu, \psi \rangle, \quad \text{as } n \to \infty,$$

for $\varphi, \psi \in C^0(Y)$. Let φ_x be the restriction of φ to Y_x and define $\widetilde{\varphi} \in C^0(X)$ by $\widetilde{\varphi}(x) = \langle \mu_x, \varphi_x \rangle$. Similarly define ψ_x and $\tilde{\psi}$. In proving (4.2) we may assume that $\varphi_x \in C^2(Y_x)$ for $x \in X$, and that $\sup_x |\varphi_x|_{C^2(Y_x)} < \infty$. Indeed, such functions are dense in $C^0(Y)$ by Lemma 1.5. Write

$$
\begin{aligned}
\langle \mu, \varphi(\psi \circ f^n) \rangle &= \int_X \langle \mu_x, \varphi_x(\psi_{x_n} \circ f_x^n) \rangle \, \mu'(x) \\
&= \int_X \left\langle \frac{1}{d^n} (f_x^n)^* \mu_{x_n}, \varphi_x(\psi_{x_n} \circ f_x^n) \right\rangle \mu'(x) \\
&= \int_X \left\langle \mu_{x_n}, \psi_{x_n} \frac{1}{d^n} (f_x^n)_* \varphi_x \right\rangle \mu'(x) \\
&= \int_X \langle \mu_{x_n}, \psi_{x_n} \widetilde{\varphi}(x) \rangle \, \mu'(x) + \int_X \left\langle \mu_{x_n}, \psi_{x_n} \left(\frac{1}{d^n} (f_x^n)_* \varphi_x - \widetilde{\varphi}(x) \right) \right\rangle \mu'(x) \\
&= \langle \mu', \widetilde{\varphi}(\tilde{\psi} \circ g^n) \rangle + \int_X \langle \mu_{x_n}, \psi_{x_n} \Phi_x \rangle \, \mu'(x),
\end{aligned}
\tag{4.3}
$$

where $\Phi_x := (1/d^n)(f_x^n)_* \varphi_x - \widetilde{\varphi}(x)$. The second equality is a consequence of Theorem 2.1 and the last line follows from the invariance of μ'. Since μ' is mixing, we know that the first term in the right-hand side of (4.3) converges to

$$\langle \mu', \widetilde{\varphi} \rangle \langle \mu', \tilde{\psi} \rangle = \langle \mu, \varphi \rangle \langle \mu, \psi \rangle.$$

To complete the proof, we will show that the integrand in the second term in the right-hand side of (4.3) converges to zero, as $n \to \infty$, uniformly in x. For this, fix $x \in X$ and pick $p, q > 1$ with $p^{-1} + q^{-1} = 1$. Write $M_1 := \sup |\psi|$, $M_2 := 2 \sup |\varphi|$ and $M_3 := \sup_x |\varphi_x|_{C^2(Y_x)}$. By Hölder's inequality we get

$$
\begin{aligned}
|\langle \mu_{x_n}, \psi_{x_n} \Phi_x \rangle| \le M_1 \langle \mu_{x_n}, |\Phi_x|^p \rangle^{1/p} &= M_1 \left(\int_0^{M_2} p t^{p-1} \mu_{x_n} \{ |\Phi_x| > t \} \, dt \right)^{1/p} \\
&\le M_1 \left(\int_0^{M_2} C p t^{p-2} d^{-n} |\varphi_x|_{C^2} \, dt \right)^{1/p} \le C_p d^{-n/p},
\end{aligned}
$$

where

$$C_p = C^{1/p} \left(\frac{p}{p-1} \right)^{1/p} M_1 M_2^{(p-1)/p} M_3^{1/p}.$$

The second inequality follows from Proposition 2.6. Since C_p does not depend on x, we see that the second term in (4.3) converges to zero, as $n \to \infty$. Thus

$$\langle \mu, \varphi(\psi \circ f^n) \rangle \to \langle \mu, \varphi \rangle \langle \mu, \psi \rangle, \quad \text{as } n \to \infty,$$

so μ is mixing for f.

To prove (ii) we suppose that μ' is ergodic for g, but not necessarily mixing. In order to show that μ is ergodic for f, we have to show that

$$\frac{1}{N}\sum_{0}^{N-1}\langle\mu,\varphi(\psi\circ f^n)\rangle\to\langle\mu,\varphi\rangle\langle\mu,\psi\rangle,\quad \text{as } N\to\infty,$$

for sufficiently regular functions φ and ψ. The proof is essentially the same as the above one; the details are left to the reader. □

5. Entropy of μ

In this section we will compute the metric entropy of the measure μ, defined by (4.1). When X is a single point the computation proves the existence of a measure of maximal entropy for (non-fibered) rational maps of $\widehat{\mathbf{C}}$. We start by recalling the definition of metric entropy. For details see [W], [R] or [Y].

Let Y be a compact metric space and ν a (completed) Borel probability measure on Y. Given a measurable partition β (not necessarily finite or countable) of Y, there exists a canonical system of conditional measures associated with β. This is a family $\{\nu_y^\beta\}$ of probability measures on Y with the following properties:

(1) for each $y\in Y$, ν_y^β is a probability measure on Y, supported on $\beta(y)$, the element of β containing y;

(2) for every measurable $E\subset Y$, $y\mapsto\nu_y^\beta(E)$ is measurable;

(3) for every measurable $E\subset Y$, $\nu(E)=\int_Y \nu_y^\beta(E)\,\nu(y)$.

For two measurable partitions α and β we define the conditional entropy of α with respect to β as

$$H_\nu(\alpha\,|\,\beta):=\int_Y -\log\nu_y^\beta(\alpha(y))\,\nu(y)$$

(this number may be infinite). If $\beta=\{Y\}$ is the trivial partition, then we write $H_\nu(\alpha):=H_\nu(\alpha|\beta)$ and we have $H_\nu(\alpha)=\int_Y -\log\nu(\alpha(y))\,\nu(y)$.

If $f\colon Y\circlearrowleft$ is continuous and leaves ν invariant, and if α is a measurable partition, then we define the entropy of f with respect to α as

$$h_\nu(f;\alpha):=H_\nu\Big(\alpha\,\Big|\,\bigvee\nolimits_{i=1}^{\infty}f^{-i}\alpha\Big).$$

Finally, the *metric entropy of* f is defined as

$$h_\nu(f):=\sup_\alpha h_\nu(f;\alpha),$$

where the supremum is taken over all measurable partitions α. In fact, it is sufficient to take the supremum over *finite* partitions.

The connection between topological and metric entropy is given by the *variational principle*, due to Goodman, Goodwyn and Dinaburg. This states that

$$h(f) = \sup_{\nu} h_{\nu}(f), \tag{5.1}$$

where the supremum is taken over all probability measures ν invariant for f. A measure ν with $h_{\nu}(f)=h(f)$ is said to be *of maximal entropy*. In general, there need not be any measure of maximal entropy, and even if there is one, it need not be unique.

We will also need to consider relative metric entropy. Let X be a compact metric space, ν' a Borel probability measure on X, and $g: X \circlearrowleft$ a continuous mapping preserving ν'. Assume that g is a *factor* of f, i.e. that there exists a continuous mapping $\pi: Y \to X$ semiconjugating f and g: $g \circ \pi = \pi \circ f$, and such that $\pi_* \nu = \nu'$. We then define the *metric entropy of f relative to g* as

$$h_{\nu}(f \mid g) := \sup_{\alpha} h_{\nu}(f \mid g; \alpha),$$

where α ranges over all measurable (or finite) partitions of Y. Here

$$h_{\nu}(f \mid g; \alpha) := H_{\nu}\Big(\alpha \,\Big|\, \bigvee_{i=1}^{\infty} f^{-i} \alpha \vee \pi^{-1}(\varepsilon_X)\Big)$$

where ε_X is the partition of X into points. See also [B] for a slightly different interpretation of relative entropy.

The connection between metric entropy and relative metric entropy is given by the Abramov–Rokhlin formula [AR]

$$h_{\nu}(f) = h_{\nu'}(g) + h_{\nu}(f \mid g); \tag{5.2}$$

see [LW, Lemma 3.1] for a proof.

Ledrappier and Walters proved the following *relativized variational principle.* In fact, they considered the more general notion of topological pressure; see Theorem 2.1 in [LW].

Theorem 5.1. *Let $g: X \circlearrowleft$ be a factor of the map $f: Y \circlearrowleft$ under the projection $\pi: Y \to X$. Let ν' be an invariant probability measure for g. Then*

$$\sup_{\nu} h_{\nu}(f \mid g) = \int_X h(f, \pi^{-1}(x))\, \nu'(x), \tag{5.3}$$

where the supremum is taken over all invariant probability measures ν for f such that $\pi_* \nu = \nu'$.

Our goal in this section is to show that the measure μ defined in (4.1) is maximal for the variational principles (5.1) and (5.3).

Theorem 5.2. *Let $f: Y\circlearrowleft$ be a rational map of degree $d\geq 2$, fibered over $g: X\circlearrowleft$, and let μ' be an invariant Borel probability measure for g. Define the measure μ by* (4.1). *Then the following holds*:

(i) $h_\mu(f|g)=\log d$;

(ii) $h_\mu(f)=h_{\mu'}(g)+\log d$.

Proof. (i) Let ε_Y be the partition of Y into points. We will compute the relative entropy $h_\mu(f|g;\varepsilon_Y)$. Write

$$\beta:=\bigvee_{i=1}^{\infty} f^{-i}(\varepsilon_Y)\vee\pi^{-1}(\varepsilon_X)=f^{-1}(\varepsilon_Y)\vee\pi^{-1}(\varepsilon_X).$$

By the construction of μ in (4.1) it follows that the conditional measures of μ with respect to the partition $\pi^{-1}(\varepsilon_X)$ are given by the measures μ_x. On the other hand, it follows from the equation $f_x^*\mu_{x_1}=d\cdot\mu_x$, and the fact that μ_x has no atoms (Theorem 2.1), that almost every element of the partition $\{f_x^{-1}f_x(z): z\in Y_x\}$ of Y_x has d elements and the conditional measures of μ_x with respect to this partition puts mass $1/d$ to each of the d points in $f_x^{-1}(f_x(z))$. Since the process of taking conditional measures is transitive, we get that for μ-almost every $z\in Y$, the element $\beta(z)$ of the partition β containing z has exactly d elements, and the conditional measure μ_z^β puts mass $1/d$ to each of these elements. We therefore have

$$h_\mu(f\,|\,g)\geq h_\mu(f\,|\,g;\varepsilon_Y)=H_\mu(\varepsilon_Y\,|\,\beta)=\int_Y -\log\mu_z^\beta(\{z\})\,\mu(z)=\log d.$$

On the other hand, Theorem 5.1 implies that

$$h_\mu(f\,|\,g)\leq\int_X h(f,Y_x)\,\mu'(x)=\log d,$$

where the last equality follows from Theorem 3.1. Thus $h_\mu(f|g)=\log d$.

(ii) This follows immediately from (i) and (5.2). □

Corollary 5.3. *If μ' is a measure of maximal entropy for g, then μ is of maximal entropy for f.*

Proof. By Theorem 3.1 we have $h(f)=h(g)+\log d$, so by Theorem 5.2(ii) it follows that μ is of maximal entropy if and only if $h_{\mu'}(g)=h(g)$, that is, if and only if μ' is of maximal entropy for g. □

6. Uniqueness of the measure of maximal entropy

In this section we will prove the converse to Theorem 5.2. This amounts to saying that the measure μ defined in (4.1) is the *unique* measure for which the suprema in the variational principles (5.1) and (5.3) are attained. In the case when X is a single point we recover the result by Lyubich [L] and Mañé [M] that a rational map of $\widehat{\mathbf{C}}$ has a unique measure of maximal entropy (see Theorem A in the introduction). The proof follows Lyubich's, although the additional difficulties arising in the fibered situation are substantial.

Theorem 6.1. *Let $f: Y \circlearrowleft$ be a rational map of degree $d \geq 2$, fibered over $g: X \circlearrowleft$. Let μ' be an invariant Borel probability measure for g. Define the measure μ by* (4.1). *Further, let ν be another invariant Borel probability measure for f such that $\pi_* \nu = \mu'$. Then the following holds*:
(i) *if $h_\nu(f \,|\, g) = \log d$, then $\nu = \mu$;*
(ii) *if $h_\nu(f) = h_\mu(f) < \infty$, then $\nu = \mu$.*

Corollary 6.2. *Let $f: Y \circlearrowleft$ and $g: X \circlearrowleft$ be as in Theorem* 6.1. *Assume that $h(g) < \infty$ and that g has a unique measure μ' of maximal entropy. Then μ, defined by* (4.1), *is the unique measure of maximal entropy for f.*

Proof. It follows from Corollary 5.3 that μ is a measure of maximal entropy for f. Suppose ν is another such measure. Write $\nu' = \pi_* \nu$. By (5.2)

$$h_{\nu'}(g) + h_\nu(f \mid g) = h_\nu(f) = h_\mu(f) = h_{\mu'}(g) + \log d.$$

But $h_{\nu'}(g) \leq h_{\mu'}(g)$ by assumption, and $h_\nu(f|g) \leq \log d$ by Theorems 5.1 and 3.1(i). This shows that the two inequalities above are in fact equalities. From the uniqueness of μ' it follows that $\nu' = \mu'$. Thus Theorem 6.1 implies that $\nu = \mu$. $\square$

The rest of this section is devoted to the proof of Theorem 6.1. We need several preliminary results, the proofs of which are, in general, deferred until the end of the section.

Lemma 6.3. *In Theorem* 6.1 *we have that* (i) *is equivalent to* (ii).

Proof. This is an immediate consequence of (5.2) and Theorem 5.2. $\square$

The next three results have nothing to do with rational maps, and hold in the following more general setting: $f: Y \circlearrowleft$ and $g: X \circlearrowleft$ are continuous mappings of compact metric spaces; $\pi: Y \to X$ is a continuous surjection that semiconjugates f to g; μ' is a g-invariant probability measure; μ and ν are f-invariant probability measures with $\pi_* \mu = \pi_* \nu = \mu'$, and with conditional measures μ_x and ν_x on the fibers Y_x of π.

Lemma 6.4. *Let f, g, Y, X, π, μ and ν be as above. Consider μ (and hence μ') as being fixed. If the inequality $h_\nu(f|g) < h_\mu(f|g)$ holds for all ergodic measures $\nu \neq \mu$, then it holds for all measures $\nu \neq \mu$.*

Lemma 6.5. *Let f, g, Y, X, π and ν be as above and assume that ν is ergodic. If $h_\nu(f|g) > 0$, then ν_x has no atoms for μ'-a.e. x.*

Lemma 6.6. *Let f, g, Y, X, π and ν be as above and assume that ν is ergodic. Let H be a compact subset of Y with $\nu(H) > 0$. Write $H_x := H \cap Y_x$ for $x \in X$. Then*

$$h_\nu(f \mid g) \leq \sup_{x \in X} h(f, H_x).$$

The next two results are specific to fibered rational maps. The first one, which is simple but crucial, says that even though a (fibered) rational map has critical points, there is an abundance of single-valued branches of inverses of high iterates.

Lemma 6.7. *Let f and g be as in Theorem 6.1, and let $n \geq l \geq 1$ and $x \in X$. Suppose that U is a conformal disk in Y_{x_n} which does not contain any critical value of $f^l_{x_{n-l}}$. Then there are at least $d^n(1-4d^{-l}(d-1))$ different single-valued branches of $(f^n_x)^{-1}$ on U.*

Lemma 6.8 below is the main technical result needed in the proof of Theorem 6.1. It says, roughly speaking, that we can find a partition of $\widehat{\mathbf{C}}$ that sufficiently distinguishes the measures μ and ν.

Lemma 6.8. *Let f, g, μ', μ and ν be as in Theorem 6.1. Assume that ν is ergodic and that $h_\nu(f|g) > 0$. Then, given any sufficiently small $\varepsilon > 0$ there exist $n \geq 1$ and a compact subset $X_3 \subset X$ with $\mu'(X_3) \geq 1-5\varepsilon$ such that for each $x \in X_3$ there is a compact subset Γ_x of Y_{x_n} such that the following holds:*

(1) *Γ_x is a finite union of smooth arcs in Y_{x_n};*

(2) *Γ_x contains all the critical values of f^n_x;*

(3) $(\mu_{x_n} + \nu_{x_n})(\Gamma_x) = 0$;

(4) *$V_x := Y_{x_n} \setminus \Gamma_x$ is a conformal disk in Y_{x_n};*

(5) *$(f^n_x)^{-1}(V_x)$ is the union of d^n disjoint conformal disks $U_{x,i}$, $i = 1, \dots, d^n$ in Y_x, and f^n_x maps each $U_{x,i}$ conformally onto V_x;*

(6) *$\mu_x(U_{x,i}) = d^{-n}$ for all i;*

(7) $\nu_x(U_{x,1}) \geq 4d^{-n}$.

Further, $\bigcup_{x \in X_3} U_{x,i}$ is relatively open in $\pi^{-1}(X_3) \subset Y$ for each i.

The proof of Theorem 6.1 also has a combinatorial part and we will need the following estimate.

Lemma 6.9. *Given numbers $d\geq 2$, $n_0\geq 1$, $\gamma\in(0,1)$ and $n\geq 1$, write*

$$A(d,n_0,\gamma,n):=\sum_{m=0}^{n_0-1}(d+1)^m+\sum_{m=n_0}^{n}\max_{\gamma m\leq j\leq m}d^{m-j}\binom{m}{j}\sum_{2d^{-1}m\leq i\leq j}\binom{j}{i}(d-1)^{j-i}.$$

Then there exists $\theta<d$ and $C>0$ such that if γ is sufficiently close to 1, *then*

$$A(d,n_0,\gamma,n)\leq C\theta^n$$

for all $n\geq 1$.

After all these preliminary results, we are now in position to prove Theorem 6.1.

Proof of Theorem 6.1. Assume that $\nu\neq\mu$, and that ν (and hence μ) is ergodic. We will show that $h_\nu(f|g)<\log d$. Assume that $h_\nu(f|g)>0$; otherwise there is nothing to prove. Choose ε so small that $1-7\varepsilon>\gamma$, with γ from Lemma 6.9. Given this ε, let n, X_3, $U_{x,i}$ and Γ_x be as in Lemma 6.8. Recall that $h_\nu(f^n|g^n)=nh_\nu(f|g)$ and $h_\mu(f^n|g^n)=nh_\mu(f|g)$. When proving Theorem 6.1 we may, and will, therefore assume that $n=1$.

Find $\delta_0>0$, a compact subset X_2 of X_3, and compact subsets V_x of $U_{x,1}$ for $x\in X_2$ such that

(1) $\mu'(X_2)\geq 1-6\varepsilon$;
(2) $\nu_x(V_x)\geq 3d^{-1}$ for $x\in X_2$;
(3) $d(V_x,\partial U_{x,1})\geq\delta_0$ for $x\in X_2$;
(4) $V:=\bigcup_{x\in X_2}V_x$ is compact.

Write $U_k=\bigcup_{x\in X_2}U_{x,k}$ for $1\leq k\leq d$.

Since μ' is ergodic we may find $n_0\geq 1$ such that the set

$$X_1:=\left\{x\in X:\frac{1}{n}\sum_{i=0}^{n-1}\chi_{X_2}(x_i)\geq 1-7\varepsilon\text{ for }n\geq n_0\right\}$$

has $\mu'(X_1)\geq 1-\varepsilon$. Further, since ν is ergodic and $\nu(V)>2d^{-1}$ we may increase n_0 so that the compact set

$$H:=\left\{y\in Y:\pi(y)\in X_1\text{ and }\frac{1}{n}\sum_{i=0}^{n-1}\chi_V(y_i)\geq 2d^{-1}\text{ for }n\geq n_0\right\}$$

has positive ν-measure.

Write $H_x:=H\cap Y_x$ for $x\in X_1$. Let θ be as in Lemma 6.9 and pick $\tilde\theta\in(\theta,d)$. We will show that $h(f,H_x)\leq\log\tilde\theta$ for $x\in X_1$. By Lemma 6.6 this implies $h_\nu(f|g)\leq\log\tilde\theta$ and therefore completes the proof.

Thus fix $x \in X_1$ with $H_x \neq \emptyset$. Let $s_m = \sum_{i=0}^{m-1} \chi_{X_2}(x_i)$ for $m \geq 1$. Note that if $m \geq n_0$, then $s_m \geq \gamma m$, by the construction of X_1.

Fix $\delta < \delta_0$ and let $n \geq n_0$. Let F_x be the (n,δ)-spanning subset of Y_x constructed in Lemma 3.3. Each element of F_x is uniquely determined by a triple (m,v,w), where $-1 \leq m \leq n$, $v \in B_m$ and $w \in Z_{m+1}$. Here $B_m = (f_x^m)^{-1}(A_m)$, where A_m is a subset of Y_{x_m} with at most Cm^3 elements. Thus $|B_m| = d^m |A_m| \leq Cm^3 d^m$. Further, Z_{m+1} is a subset of $Y_{x_{m+1}}$ with at most Cm elements. The triple (m,v,w) determines $z(m,v,w)$ in the following way:

$$\begin{aligned} d(f_x^i z, f_x^i v) &< \tfrac{1}{2}\delta \quad \text{for } 0 \leq i \leq m, \\ d(f_x^i z, f_{x_{m+1}}^{i-m-1} w) &< \tfrac{1}{2}\delta \quad \text{for } m < i \leq n. \end{aligned}$$

Let Ω_x be the dense subset of Y_x defined by

$$\Omega_x = \{ y \in Y_x : y_{i+1} \notin \Gamma_{x_i} \text{ whenever } i \geq 0 \text{ and } x_i \in X_2 \}.$$

We may assume that $B_m \subset \Omega_x$ for each m. Thus $v \in \Omega_x$ for each triple (m,v,w).

Pick a minimal subset E_x of F_x which (n,δ)-spans H_x. Let $z(m,v,w) \in E_x$. By the minimality of E_x there exists $y \in H_x$ such that $d(f^i y, f^i v) < \delta$ for $0 \leq i < m$. By the choice of δ_0 we therefore have

$$f^i y \in V \quad \Longrightarrow \quad f^i v \in U_1$$

for $0 \leq i < m$. Thus, if $m \geq n_0$, then

$$\frac{1}{m} \sum_{i=0}^{m-1} \chi_{U_1}(f^i v) \geq \frac{1}{m} \sum_{i=0}^{m-1} \chi_V(f^i y) \geq 2d^{-1}. \tag{6.1}$$

To each $v \in B_m$ we assign a sequence

$$\alpha(v) = (\alpha_0(v), \alpha_1(v), \dots, \alpha_{m-1}(v)) \in \{0, \dots, d\}^m,$$

where

$$\alpha_i(v) = \begin{cases} 0, & \text{if } x_i \notin X_2, \\ k, & \text{if } x_i \in X_2 \text{ and } f^i v \in U_k. \end{cases}$$

This makes sense, because $v \in \Omega_x$.

For $0 \leq j \leq m$ let $D_{m,j}$ be the set of sequences

$$(\alpha_0, \dots, \alpha_{m-1}) \in \{0, \dots, d\}^m$$

such that $\alpha_k \neq 0$ for exactly j indices k, and $\alpha_k = 1$ for at least $2m/d$ indices k. Then

$$|D_{m,j}| = \binom{m}{j} \sum_{i=2m/d}^{j} \binom{j}{i} (d-1)^{j-1}. \tag{6.2}$$

Note that if $z(m,v,w) \in E_x$, then $\alpha(v) \in D_{m,s_m}$ by (6.1). Further, we claim that if $m \geq n_0$, then given $\alpha \in D_{m,s_m}$ and $u \in A_m$ there exist exactly d^{m-s_m} points $v \in B_m$ such that $f_x^m v = u$ and $\alpha(v) = \alpha$. To prove this claim, recall that if $i \geq 0$ and $x_i \in X_2$, then we defined above a conformal disk V_{x_i} in $Y_{x_{i+1}}$ containing all the d^{m-i-1} values of $(f_{x_i}^{m-i-1})^{-1}(u)$, but containing no critical values of f_{x_i}. If $0 \leq i < m$ and $x_i \notin X_2$, then let V_{x_i} be any conformal disk in $Y_{x_{i+1}}$ having the same two properties. Any point $v \in (f_x^m)^{-1}(u)$ can then be written as $v = g_0 \circ \ldots \circ g_{m-1}(u)$, where g_i is a single-valued branch of $f_{x_i}^{-1}$ defined on V_{x_i}. Now suppose $\alpha(v) = \alpha$, where $\alpha = (\alpha_0, \ldots, \alpha_{m-1}) \in D_{m,s_m}$. If $x_i \in X_2$, then the branch g_i is uniquely determined by the number α_i. If $x_i \neq X_2$, then there are exactly d choices for g_i. Thus there are exactly d^{m-s_m} points $v \in B_m$ such that $f_x^m v = u$ and $\alpha(v) = \alpha$, which proves the claim.

It follows that

$$\begin{aligned}
|E_x| &\leq \sum_{m=0}^{n_0-1} d^m |A_m|\,|Z_{m+1}| + \sum_{m=n_0}^{n} d^{m-s_m} |D_{m,s_m}|\,|A_m|\,|Z_{m+1}| \\
&\leq \left(\sum_{m=0}^{n_0-1} d^m + \sum_{m=n_0}^{n} d^{m-s_m} |D_{m,s_m}| \right) Cn^3 Cn \\
&\leq Cn^4 A(d, n_0, \gamma, n) \leq Cn^4 \theta^n \leq C\tilde{\theta}^n,
\end{aligned}$$

by (6.2) and Lemma 6.9. We have constructed an (n,δ)-spanning subset of H_x with at most $C\tilde{\theta}^n$ elements. Thus $h(f, H_x) \leq \log \tilde{\theta}$, and we are done. □

In the rest of the section we give the proofs of the various results needed to prove Theorem 6.1. We start with the general results on fibered mappings (Lemmas 6.4, 6.5 and 6.6).

Proof of Lemma 6.4. We will use the decomposition of invariant measures into ergodic components. Let $M(Y,f)$ be the set of f-invariant probability measures on Y and let $E(Y,f)$ be the subset of ergodic measures. Then $M(Y,f)$ and $E(Y,f)$ are compact in the weak topology. In fact, they are metrizable. Analogously we define $M(X,g)$ and $E(X,g)$. The map $\pi_*: M(Y,f) \to M(X,g)$ induced from $\pi: Y \to X$ maps $E(Y,f)$ into $E(X,g)$. Given $\mu' \in M(X,g)$ define a measure $\tau(\mu') = \mu$ on Y by (4.1). It follows from Proposition 4.1 and Theorem 4.2 that τ maps $M(X,g)$ to $M(Y,f)$ and $E(X,g)$ to $E(Y,f)$. Further, $\pi_* \circ \tau = \mathrm{id}$.

The ergodic (Rokhlin) decomposition tells us that given $\mu'\in M(X,g)$ there exists a unique probability measure $\varrho_{\mu'}$ on $E(X,g)$ such that

$$\mu'=\int_{E(X,g)}\sigma'\,\varrho_{\mu'}(\sigma').$$

The map τ commutes with this decomposition in the sense that

$$\mu=\tau\mu'=\int_{E(X,g)}\tau\sigma'\,\varrho_{\mu'}(\sigma').$$

Now let $\nu\in M(Y,f)$ be a measure with $\pi_*\nu=\mu'$ and write

$$\nu=\int_{E(Y,f)}\sigma\,\varrho_\nu(\sigma), \tag{6.3}$$

where ϱ_ν is a probability measure on $E(Y,f)$. We have $(\pi_*)_*\varrho_\nu=\varrho_{\mu'}$.

The existence of τ above shows that $\pi_*\colon E(Y,f)\to E(X,g)$ is a continuous surjection, the fibers of which constitute a measurable partition of $E(Y,f)$. Let $\{\varrho_{\nu,\sigma'}\}_{\sigma'\in E(X,g)}$ be the conditional measures of ϱ_ν with respect to this partition. Thus

$$\nu=\int_{E(X,g)}\varrho_{\mu'}(\sigma')\int_{\pi_*^{-1}\{\sigma'\}}\sigma\,\varrho_{\nu,\sigma'}(\sigma).$$

We now use the fact that the relative metric entropy $h_\nu(f\,|\,g)$ commutes with the ergodic decomposition (see [LW, Lemma 3.2(iii)]),

$$h_\nu(f\mid g)=\int_{E(Y,f)}h_\sigma(f\mid g)\,\varrho_\nu(\sigma)=\int_{E(X,g)}\varrho_{\mu'}(\sigma')\int_{\pi_*^{-1}\{\sigma'\}}h_\sigma(f\mid g)\,\varrho_{\nu,\sigma'}(\sigma).$$

Suppose that $h_\nu(f|g)=h_\mu(f|g)=\log d$. By Theorem 3.1(i) we have $h_\sigma(f|g)\le\log d$ for all $\sigma\in M(Y,f)$. Thus $h_\sigma(f|g)=\log d$ for $\varrho_{\mu'}$-a.e. σ' and $\varrho_{\nu,\sigma'}$-a.e. σ. But then the assumptions of the lemma imply that $\sigma=\tau\sigma'$ for $\varrho_{\mu'}$-a.e. σ' and $\varrho_{\nu,\sigma'}$-a.e. σ. Thus

$$\nu=\int_{E(X,g)}(\tau\sigma')\,\varrho_{\mu'}(\sigma')=\mu,$$

which completes the proof. □

In order to prove Lemma 6.5 we use the following result. It is a relativized version of the Shannon–McMillan–Breiman theorem. For the proof (in a slightly different situation), see Theorem 4.2 in [B].

Lemma 6.10. *Let f, g, Y, X, π and ν be as in Lemma 6.5. Let α be a finite measurable partition of Y. Then*

$$-\frac{1}{n}\log \nu_{\pi(y)}(\alpha^n(y)) \to h_\nu(f \mid g;\alpha)$$

for ν-almost every y. Here $\alpha^n(y)$ denotes the element of the partition α^n containing y.

Proof of Lemma 6.5. Assume that $h_\nu(f|g)>0$ and find a finite partition α such that $h_\nu(f|g;\alpha)>0$. Let Y_0 be the set of $y\in Y$ such that y is an atom for $\nu_{\pi(y)}$. If $y\in Y_0$ and $\alpha^n(y)$ is the element of α^n containing y, then

$$0\le -\frac{1}{n}\log \nu_{\pi(y)}(\alpha^n(y)) \le -\frac{1}{n}\log \nu_{\pi(y)}(\{y\}) \to 0, \quad \text{as } n\to\infty.$$

By Lemma 6.10 and the assumption that $h(f|g;\alpha)>0$ it follows that $\nu(Y_0)=0$. Thus ν_x has no atoms for μ'-a.e. x. □

Proof of Lemma 6.6. Let $\alpha=\{A_1,\dots,A_s\}$ be a finite partition of Y. Pick compact subsets $B_i\subset A_i$ and let $B_0:=Y\setminus\bigcup_{i=1}^s B_i$. We may choose B_i close enough to A_i so that $H_\nu(\alpha|\beta)\le 1$, where $\beta=\{B_0,B_1,\dots,B_s\}$. Therefore

$$h_\nu(f \mid g;\alpha) \le h_\nu(f \mid g;\beta) + H_\nu(\alpha \mid \beta\vee\pi^{-1}(\varepsilon_X)) \le h_\nu(f \mid g;\beta)+1.$$

By Lemma 6.10 and Egorov's theorem there exists a compact subset H' of H with $\nu(H')>0$ such that

$$\frac{1}{n}\log \nu_{\pi(y)}(\beta^n(y)) \to -h_\nu(f \mid g;\beta)$$

uniformly on H'. Here $\beta^n(y)$ denotes the element of β^n containing y. Write $H'_x=H'\cap Y_x$ for $x\in X$ and fix x such that $\nu_x(H'_x)>0$.

Pick $\delta>0$ small. For large n and $y\in H'_x$ we have

$$\nu_x(\beta^n(y)) \ge e^{-n(h_\nu(f|g;\beta)-\delta)}.$$

Let E_x be a minimal (n,δ)-spanning subset of H'_x. For each $C\in\beta^n$ with $C\cap H'_x\ne\emptyset$ we associate a point $z(C)\in E_x$ with $B_n(z,\delta)\cap C\ne\emptyset$, where

$$B_n(z,\delta)=\{w\in Y_x : d(f^i w, f^i z)<\delta \text{ for } 0\le i<n\}.$$

Suppose that $z(C)=z(C')$ for some elements $C,C'\in\beta^n$ such that $C\cap H'_x\ne\emptyset$ and $C'\cap H'_x\ne\emptyset$. Then there exist $w\in C\cap H'_x$, $w'\in C'\cap H'_x$ such that $d(f^i w, f^i w')<2\delta$ for

$0 \le i < n$. If δ is small enough, then this implies that for each i, $f^i w$ and $f^i w'$ belong to the same element of the covering $\{B_0 \cup B_1, \dots, B_0 \cup B_s\}$ of Y. Hence, for each i there exists $j = j(i)$ such that $f^i w, f^i w' \in B_0 \cup B_j$. It follows that there are at most 2^n different elements $C \in \beta^n$ associated with a given point in E_x. Thus

$$0 < \nu_x(H'_x) \le \sum_{\substack{C \in \beta^n \\ C \cap H'_x \ne \emptyset}} \nu_x(C) \le 2^n |E_x| e^{-n(h_\nu(f|g;\beta)-\delta)} = 2^n r_n(\delta, H'_x) e^{-n(h_\nu(f|g;\beta)-\delta)}.$$

This implies that

$$\frac{1}{n} \log r_n(\delta, H_x) \ge \frac{1}{n} \log r_n(\delta, H'_x) \ge h_\nu(f \mid g; \beta) - \delta - \log 2 + \frac{1}{n} \log \nu_x(H'_x).$$

By letting $n \to \infty$ and $\delta \to 0$ we obtain

$$h(f, H_x) \ge h_\nu(f \mid g; \beta) - \log 2 \ge h_\nu(f \mid g; \alpha) - 1 - \log 2.$$

After replacing f by f^n we obtain

$$nh(f, H_x) \ge nh_\nu(f \mid g; \alpha) - 1 - \log 2.$$

By dividing this inequality by n and letting $n \to \infty$ we see that

$$h_\nu(f \mid g; \alpha) \le h(f, H_x).$$

Since α was an arbitrary finite partition we have shown that

$$h_\nu(f \mid g) \le \sup_{x \in X} h(f, H_x). \quad \square$$

We next prove Lemmas 6.7 and 6.8 that are specific to fibered rational maps.

Proof of Lemma 6.7. For $1 \le m \le n$ let σ_m be the number of single-valued branches of $(f^m_{x_{n-m}})^{-1}$ on U. We will show by induction on m that

$$\sigma_m \ge d^m - 2(d-1) \sum_{i=1}^{m-l} d^i, \tag{6.4}$$

which clearly implies the assertion of the lemma when $m = n$.

If $1 \le m \le l$, then U contains no critical values of $f^m_{x_{n-m}}$, so $\sigma_m = d^m$ and (6.4) is trivial. For the inductive step, consider the σ_{m-1} single-valued branches of $(f^{m-1}_{x_{n-m+1}})^{-1}$ defined on U. At most $2(d-1)$ of the images of these branches can contain a critical value of $f_{x_{n-m}}$. If we compose the other branches with all possible branches of $f^{-1}_{x_{n-m}}$, then we see that

$$\sigma_m \geq d(\sigma_{m-1}-2(d-1)) \geq d^m-2(d-1)\sum_{i=1}^{m-l} d^i,$$

where the last inequality follows from (6.4) for $m-1$. □

Proof of Lemma 6.8. Assume that $\varepsilon<\frac{1}{100}$. Pick l so large that $4(d-1)d^{-l}<\varepsilon$.

The assumptions imply that μ and ν are mutually singular. We may therefore find disjoint compact sets L_μ and L_ν in Y such that $\mu L_\mu\geq 1-\frac{1}{2}\varepsilon^2$ and $\nu L_\nu\geq 1-\frac{1}{2}\varepsilon^2$. Write $L_{\mu,x}=L_\mu\cap Y_x$ and $L_{\nu,x}=L_\nu\cap Y_x$ for $x\in X$. We may find a compact subset X_7 of X with $\mu'(X_7)\geq 1-\varepsilon$ such that $\mu_x(L_{\mu,x})\geq 1-\varepsilon$ and $\nu_x(L_{\nu,x})\geq 1-\varepsilon$ for $x\in X_7$. Pick $\varkappa>0$ such that $\varkappa<d(L_{\mu,x},L_{\nu,x})$ for all $x\in X_7$.

We will find a compact subset X_6 of X_7 with $\mu'(X_6)\geq 1-3\varepsilon$, such that the following holds: for each $x\in X_6$ there exists a compact subset Γ'_x of Y_{x_l} such that

(1) Γ'_x is a finite union of smooth arcs in Y_{x_l};

(2) Γ'_x contains all the critical values of f^l_x;

(3) $(\mu_{x_l}+\nu_{x_l})(\Gamma'_x)=0$;

(4) $V'_x:=Y_{x_l}\setminus\Gamma'_x$ is a conformal disk in Y_{x_l}.

Further, the arcs in Γ'_x depend continuously on x.

To do this, first write $X_7=X''_1\cup...\cup X''_m$, where X''_i are disjoint Borel sets such that Y is trivial over $g^l(X''_i)$, and such that the critical values of f^l_x depend continuously on x for $x\in X''_i$ in the sense that there exist continuous functions $s_j\colon X''_i\to Y$ such that $s_j(x)\in Y_{x_l}$, $s_{j_1}(x)\neq s_{j_2}(x)$ for $j_1\neq j_2$ and such that $\{s_j(x)\}$ is the set of critical values of f^l_x.

Pick $X'_i\subset X''_i$ compact for $i=1,\dots,m$ such that $\sum_{i=1}^m \mu'(X''_i\setminus X'_i)<\varepsilon$. It suffices to construct Γ'_x for $x\in X'_i$ for each i individually. To simplify notation we fix i and write X' instead of X'_i.

Fix $x^0\in X'$ and write $x^0_l=g^l(x^0)$. Let $\Phi\colon g^l(X')\times\widehat{\mathbf{C}}\to\pi^{-1}(g^l(X'))\subset Y$ be a trivialization of $\pi\colon Y\to X$. Using Φ, it makes sense to talk about spherical arcs in Y_x for $x\in g^l(X')$. Let $\{s_j(x)\}$ be the critical values of f^l_x as above. Let Γ'' be a finite union of spherical arcs in $Y_{x^0_l}$ of length <1, the endpoints of which contain all the critical values of $f^l_{x^0}$. Assume that the arcs in Γ'' only intersect at endpoints, and that $Y_{x^0_l}\setminus\Gamma''$ is connected and simply connected, i.e. a conformal disk. We may find a neighborhood Ω of x^0 in X' and a finite number of continuous functions $\alpha_k,\beta_k\colon\Omega\to Y$ such that

(1) $\alpha_k(x),\beta_k(x)\in Y_{x_l}$ for all k and all $x\in\Omega$;

(2) $\bigcup_k\{\alpha_k(x),\beta_k(x)\}$ contains all the critical values of f^l_x for all $x\in\Omega$;

(3) for all j and k, each of the sets $\{x\in\Omega:\alpha_j(x)=\alpha_k(x)\}$, $\{x\in\Omega:\alpha_j(x)=\beta_k(x)\}$ and $\{x\in\Omega:\beta_j(x)=\beta_k(x)\}$ is either empty or all of Ω;

(4) the set of arcs $[\alpha_k(x^0),\beta_k(x^0)]$ coincides with the set of arcs in Γ''.

Let $\psi_k:\mathbf{D}\to\widehat{\mathbf{C}}$ be a univalent meromorphic function such that $\Phi(x_l^0,\psi_k(0))$ is the midpoint on the arc $[\alpha_k(x^0),\beta_k(x^0)]$. For $\lambda\in\mathbf{D}$ and $x\in\Omega$ let $\gamma_k(x,\lambda)=\Phi(x_l,\psi_k(\lambda))$. Let $\Gamma'_x(\lambda)$ be the union of all arcs $[\alpha_k(x),\gamma_k(x,\lambda)]$ and $[\gamma_k(x,\lambda),\beta_k(x)]$. Note that $\Gamma'_{x^0}(0)=\Gamma''$. After shrinking Ω if necessary, we have that if $x\in\Omega$, then $Y_{x_l}\setminus\Gamma'_x(\lambda)$ is a conformal disk in Y_{x_l} for all sufficiently small λ. Write $\Gamma'(\lambda)=\bigcup_{x\in\Omega}\Gamma'_x(\lambda)$. Since ν_{x_l} and μ_{x_l} have no atoms (Lemma 6.5 and Theorem 2.1) and the sets $\Gamma'(\lambda_1)$ and $\Gamma'(\lambda_2)$ are essentially disjoint if $\lambda_1\neq\lambda_2$, we may find a small $\lambda\in\mathbf{D}$ such that $(\mu+\nu)(\Gamma'(\lambda))=0$. Fix such a λ. Then there exists a subset $\Omega'\subset\Omega$ of full μ'-measure such that $(\mu_{x_l}+\nu_{x_l})(\Gamma'_x(\lambda))=0$ for $x\in\Omega'$. Write $\Gamma'_x=\Gamma'_x(\lambda)$.

Returning to our previous notation, we have constructed Γ'_x for x in a subset of full measure of a neighborhood of any point $x^0\in X'_i$ for any i. It is now easy to find disjoint compact subsets $X'_{i,1},\dots,X'_{i,k(i)}$ of X'_i such that $\sum_{j=1}^{k(i)}\mu'(X'_{i,j})\geq(1-\varepsilon)\mu'(X'_i)$ and such that Γ'_x can be constructed as above on each $X'_{i,j}$. Let X_6 be the union of all the sets $X'_{i,j}$. Then $\mu'(X_6)\geq 1-3\varepsilon$ and Γ'_x exists for $x\in X_6$.

For $x\in X_6$ and $\varepsilon'>0$, let $\widetilde{\Gamma}_x$ be the ε'-fattening of Γ'_x, i.e.

$$\widetilde{\Gamma}_x=\{z\in Y_{x_l}:d(z,\Gamma'_x)<\varepsilon'\}.$$

Since $x\mapsto\Gamma'_x$ is continuous, we may choose ε' so small that $Y_{x_l}\setminus\overline{\widetilde{\Gamma}}_x$ is a conformal disk and that $\widetilde{\Gamma}_x\setminus\Gamma'_x$ is a conformal annulus for all $x\in X_6$. Further, we may find a compact subset $X_5\subset X_6$ with $\mu'(X_5)\geq 1-4\varepsilon$ such that for $x\in X_5$,

(1) $(\mu_{x_l}+\nu_{x_l})(\overline{\widetilde{\Gamma}}_x)\leq\varepsilon$;

(2) the modulus of $\widetilde{\Gamma}_x\setminus\Gamma'_x$ is bounded below by a positive constant.

By the Koebe distortion theorem, the latter property implies that if $\psi:Y_{x_l}\setminus\Gamma'_x\to\widehat{\mathbf{C}}$ is a univalent meromorphic function, then

$$\operatorname{diam}(\psi(Y_{x_l}\setminus\overline{\widetilde{\Gamma}}_x))\leq C\sqrt{\operatorname{area}(\psi(Y_{x_l}\setminus\overline{\widetilde{\Gamma}}_x))} \tag{6.5}$$

for some constant $C>0$.

Let n be a large number (how large will be seen later). Write $X_4=g^{-(n-l)}X_5$. Then X_4 is compact and $\mu'(X_4)\geq 1-4\varepsilon$.

Using essentially the same procedure as when we constructed Γ'_x we may find a compact subset $X_3\subset X_4$ with $\mu'(X_3)\geq 1-5\varepsilon$ such that the following holds: for each $x\in X_3$ there exists a compact subset $\Gamma_x\subset Y_{x_n}$, depending continuously on x, such that

(1) Γ_x is a finite union of smooth arcs in Y_{x_n};

(2) $Y_{x_n}\setminus\Gamma_x$ is a conformal disk;
(3) $\Gamma_x\supset\Gamma'_{x_{n-l}}$;
(4) $(\mu_{x_n}+\nu_{x_n})(\Gamma_x)=0$.
For $x\in X_3$ define

$$\begin{aligned} V_x &= Y_{x_n}\setminus\Gamma_x,\\ V'_x &= Y_{x_n}\setminus\Gamma'_{x_{n-l}},\\ \widetilde{V}_x &= Y_{x_n}\setminus\overline{\overline{\Gamma}}_{x_{n-l}}. \end{aligned}$$

Then V_x, V'_x and $\widetilde{V}_x$ are all conformal disks in Y_{x_n}. Clearly $V_x\subset V'_x$ and $\widetilde{V}_x\subset V'_x$. Further,

$$\mu_{x_n}(V_x)=\nu_{x_n}(V_x)=\mu_{x_n}(V'_x)=\nu_{x_n}(V'_x)=1$$

and

$$\mu_{x_n}(\widetilde{V}_x)\geq 1-\varepsilon,\quad \nu_{x_n}(\widetilde{V}_x)\geq 1-\varepsilon.$$

Write

$$\begin{aligned} (f_x^n)^{-1}V_x &= \bigcup_{i\in I} U_{x,i},\\ (f_x^n)^{-1}V'_x &= \bigcup_{j\in J} U'_{x,j},\\ (f_x^n)^{-1}\widetilde{V}_x &= \bigcup_{k\in K} \widetilde{U}_{x,k}, \end{aligned}$$

where $U_{x,i}$, $U'_{x,j}$ and $\widetilde{U}_{x,k}$ are conformal disks in Y_x, and the three unions are disjoint. Fix $x\in X_3$. We will show that there exists $i\in I$ such that

$$\nu_x(U_{x,i})\geq 4\mu(U_{x,i})=4d^{-n}.$$

For each $k\in K$ there exists a unique $j=j(k)\in J$ such that $\widetilde{U}_{x,k}\subset U'_{x,j}$. Let K_1 be the set of $k\in K$ such that f_x^n is univalent on $U_{x,j(k)}$. Recall that V'_x contains no critical values of $f^l_{x_{n-l}}$. Thus $|K_1|\geq d^n(1-4d^{-l}(d-1))\geq d^n(1-\varepsilon)$ by Lemma 6.7.

If $k\in K_1$, then we may apply (6.5) to the branch of $(f_x^n)^{-1}$ mapping V'_x to $U'_{x,k}$. Assume that n is so large that

$$C^2\operatorname{area}(\widehat{\mathbf{C}})\leq\varepsilon\varkappa^2 d^n,$$

where $\varkappa<d(L_{\mu,x},L_{\nu,x})$ was chosen above. Let K_2 be the set of $k\in K_1$ such that $\operatorname{diam}(\widetilde{U}_{x,k})<\varkappa$. Then $|K_2|\geq d^n(1-2\varepsilon)$ by the above estimate. Let K_3 be the set of $k\in K_2$ such that $\widetilde{U}_{x,k}\cap L_{\mu,x}\neq\emptyset$. Since

$$\mu_x(\widetilde{U}_{x,k})=d^{-n}\mu_{x_n}(\widetilde{V}_x)\leq d^{-n}$$

for every $k\in K_1$, and since $\mu_x(L_{\mu,x})\geq 1-\varepsilon$, we have $|K_3|\geq d^n(1-3\varepsilon)$. Note that if $k\in K_3$, then $\widetilde{U}_{x,k}\cap L_{\nu,x}=\emptyset$.

Given $i\in I$ there exists a unique $j=j(i)\in J$ such that $U_{x,i}\subset U'_{x,j}$. If f_x^n is univalent on $U'_{x,j(i)}$, then there also exists a unique $k=k(i)\in K_1$ such that $\widetilde{U}_{x,k}\subset U'_{x,j}$. Let I_1 be the set of $i\in I$ such that f_x^n is univalent on $U'_{x,j(i)}$ and such that $k(i)\in K_3$. We have

$$\mu_x\Big(\bigcup_{i\in I_1} U_{x,i}\Big)=\mu_x\Big(\bigcup_{i\in I_1} U'_{x,j(i)}\Big)\geq \mu_x\Big(\bigcup_{k\in K_3} \widetilde{U}_{x,k}\Big)\geq 1-3\varepsilon$$

and

$$\nu_x\Big(\bigcup_{i\in I_1} U_{x,i}\Big)\leq \nu_x((f_x^n)^{-1}(V_x\setminus\widetilde{V}_x))+\nu_x\Big(\bigcup_{k\in K_3} \widetilde{U}_{x,k}\Big)\leq \varepsilon+\varepsilon,$$

where we have used the fact that $\widetilde{U}_{x,k}\cap L_{\nu,x}=\emptyset$ for $x\in K_3$.

Let $I_2=I\setminus I_1$. Then

$$\mu_x\Big(\bigcup_{i\in I_2} U_{x,i}\Big)\leq 3\varepsilon,$$

$$\nu_x\Big(\bigcup_{i\in I_2} U_{x,i}\Big)\geq 1-2\varepsilon.$$

Thus there exists $i\in I_2$ such that

$$\nu_x(U_{x,i})\geq \frac{1-2\varepsilon}{3\varepsilon}\mu_x(U_{x,i})\geq 4d^{-n}. \tag{6.6}$$

By relabeling the disks $U_{x,i}$ we may assume that $I=\{1,\dots,d^n\}$, that (6.6) holds for $i=1$, and that $\bigcup_{x\in X_3} U_{x,i}$ is relatively open in $\pi^{-1}(X_3)$ for $1\leq i\leq d^n$. This completes the proof of Lemma 6.8. $\square$

Finally we will prove the estimate in Lemma 6.9. For this we need the following elementary result.

Lemma 6.11. *Let φ be the continuous function on $\{(x,y):0\leq y\leq x\leq 1\}$ defined by*

$$\varphi(x,y)=(1-x)\log d+(x-y)\log(d-1)-(1-x)\log(1-x)-(x-y)\log(x-y)-y\log y$$

for some $d\geq 2$. Then there exist $\theta'<d$ and $\gamma<1$ such that $\varphi\leq\log\theta'$ on the set

$$\{(x,y):\gamma\leq x\leq 1,\ 2/d\leq y\leq x\}.$$

Proof. We first consider the restriction of φ to the line $x=1$, i.e.

$$\varphi(1,y)=(1-y)\log(d-1)-(1-y)\log(1-y)-y\log y.$$

It is easy to verify that $y\mapsto\varphi(1,y)$ is decreasing on the interval $1/d\leq y\leq 1$ and that $\varphi(1,1/d)=\log d$. If we let $\theta''=\exp\varphi(1,2/d)$, then $\theta''<d$ and $\varphi(1,y)\leq\log\theta''$ for $2/d\leq y\leq 1$. By continuity it follows that if $\theta''<\theta'<d$ and γ is sufficiently close to 1, then $\varphi(x,y)\leq\log\theta'$ when $\gamma\leq x\leq 1$ and $2/d\leq y\leq x$. $\square$

Proof of Lemma 6.9. Throughout the proof, C will denote different positive constants, depending on d, n_0 and γ but not on n. Let φ, θ' and γ be as in Lemma 6.11 and let $\theta\in(\theta',d)$.

The first sum in $A(d,n_0,\gamma,n)$ does not depend on n. To prove the lemma it is therefore sufficient to show that

$$d^{m-j}(d-1)^{j-i}\binom{m}{j}\binom{j}{i}\leq C\theta^n \tag{6.7}$$

whenever $n_0\leq m\leq n$, $\gamma m\leq j\leq m$ and $2d^{-1}m\leq i\leq j$, because Lemma 6.9 then follows with a slightly larger θ.

By Stirling's formula there exists $r>0$ such that

$$\binom{m}{j}\leq C\frac{m^m}{j^j(m-j)^{m-j}}\quad\text{and}\quad\binom{j}{i}\leq C\frac{j^j}{i^i(j-i)^{j-i}} \tag{6.8}$$

for $r\leq j\leq m-r$ and $r\leq i\leq j-r$.

To prove (6.7) we consider four cases. The first case is when $m-r\leq j\leq m$ and $j-r\leq i\leq j$. Then

$$d^{m-j}(d-1)^{j-i}\binom{m}{j}\binom{j}{i}\leq Cm^{2r}\leq C\theta^n.$$

The second case is when $\gamma m\leq j\leq m-r$ and $j-r\leq i\leq j$. Then $\gamma\leq j/m\leq 1$, so by (6.8) and Lemma 6.11 we have

$$\begin{aligned}d^{m-j}(d-1)^{j-i}\binom{m}{j}\binom{j}{i}&\leq Cm^r d^{m-j}\frac{m^m}{j^j(m-j)^{m-j}}\\&=Cm^r\exp\left(m\varphi\left(\frac{j}{m},\frac{j}{m}\right)\right)\leq Cm^r(\theta')^m\leq C\theta^n.\end{aligned}$$

The third case is when $m-r\leq j\leq m$ and $2d^{-1}j\leq i\leq j-r$. Then $2/d\leq i/j\leq 1$, so by (6.8) and Lemma 6.11 we have

$$\begin{aligned}d^{m-j}(d-1)^{j-i}\binom{m}{j}\binom{j}{i}&\leq Cm^r(d-1)^{j-i}\frac{j^j}{i^i(j-i)^{j-i}}\\&=Cm^r\exp\left(m\varphi\left(1,\frac{i}{j}\right)\right)\leq Cm^r(\theta')^m\leq C\theta^n.\end{aligned}$$

The last case, finally, is when $\gamma m \leq j \leq m-r$ and $2d^{-1}j \leq i \leq j-r$. Then we have $\gamma \leq j/m \leq 1$ and $2/d \leq i/m \leq 1$, so by (6.8) and Lemma 6.11 we obtain

$$\begin{aligned} d^{m-j}(d-1)^{j-i}\binom{m}{j}\binom{j}{i} &\leq Cd^{m-j}(d-1)^{j-i}\frac{m^m}{j^j(m-j)^{m-j}}\frac{j^j}{i^i(j-i)^{j-i}} \\ &= C\exp\biggl(m\varphi\biggl(\frac{i}{m},\frac{j}{m}\biggr)\biggr) \leq C(\theta')^m \leq C\theta^n. \end{aligned}$$

This completes the proof of Lemma 6.9. □

7. Applications to complex surfaces

In this section we apply our techniques to dynamics on ruled surfaces. Our main result is that, with some restrictions, a holomorphic selfmap of a ruled surface has a unique measure of maximal entropy. The same conclusion will also be drawn for certain holomorphic mappings of $\mathbf{P}^2$.

For us, a *ruled surface* is a smooth projective complex surface Y which is a holomorphic $\mathbf{P}^1$-bundle over a compact Riemann surface X. This means that there is a holomorphic projection $\pi: Y \to X$ such that $\pi^{-1}(x) \simeq \mathbf{P}^1 \simeq \widehat{\mathbf{C}}$ for every $x \in X$ and such that every $x \in X$ has an open neighborhood U with $\pi^{-1}(U) \simeq U \times \widehat{\mathbf{C}}$.

The following result by Dabija tells us that selfmaps of ruled surfaces can be viewed as fibered rational maps in the sense of Section 2.

Proposition 7.1. (Proposition 7.1 in [D]) *Let Y be a ruled surface over X and let $f: Y\circlearrowleft$ be a holomorphic mapping.*

(i) *If $Y \neq \mathbf{P}^1 \times \mathbf{P}^1$, then there exists a holomorphic mapping $g: X\circlearrowleft$ such that π semiconjugates f to g: $g \circ \pi = \pi \circ f$.*

(ii) *If $Y = \mathbf{P}^1 \times \mathbf{P}^1$, then the same conclusion holds for $f^2 = f \circ f$ instead of f.*

Remark 7.2. If $Y = \mathbf{P}^1 \times \mathbf{P}^1$, then f may be of the form $f(z,w) = (p(w), q(z))$, where p and q are rational functions.

By Proposition 7.1 we may apply the results of the preceding sections to study the dynamics of selfmaps of ruled surfaces.

Theorem 7.3. *Let Y be a ruled surface over X and let f be a holomorphic mapping of Y which fibers over a holomorphic map $g: X\circlearrowleft$. Assume that the topological degrees δ_f and δ_g of f and g satisfy $1 < \delta_g < \delta_f$. Then $h(f) = \log \delta_f$ and f has a unique measure of maximal entropy.*

Proof. We first consider the dynamics of g. For this we use the classification of compact Riemann surfaces. First note that X cannot be hyperbolic, because then g

would have to be an automorphism or a constant mapping, contradicting $\delta_g>1$. If X is a torus, then since $\delta_g>1$, $g: X \circlearrowleft$ is an expanding linear map. Thus $h(g)=\log\delta_g$ and g has a unique measure μ' of maximal entropy (this is just the pushforward of Lebesgue measure on $\mathbf{C}$ under the universal covering map). If X is the Riemann sphere $\widehat{\mathbf{C}}$, then g is a rational map of degree $\delta_g>1$. By results for (non-fibered) rational maps (see Theorem A in the introduction), $h(g)=\log\delta_g$ and g has a unique measure μ' of maximal entropy.

Now f is a rational map fibered over g of degree $d:=\delta_f/\delta_g>1$. Define the measure μ by (4.1). It then follows from Corollary 6.2 that μ is the unique measure of maximal entropy for f. Finally

$$h(f)=h(g)+\log d=\log\delta_g+\log(\delta_f/\delta_g)=\log\delta_f,$$

by Theorem 3.1. This completes the proof. □

We now turn to holomorphic mappings of $\mathbf{P}^2$. It is known that such mappings have a measure of maximal entropy [FS3]; this measure can be quite explicitly described. However, it is an open problem whether the measure is the unique one with maximal entropy.

Unfortunately, $\mathbf{P}^2$ is not a ruled surface, so we cannot apply Theorem 7.3 to solve this problem. What we will do here is to restrict our attention to a certain class of selfmaps of $\mathbf{P}^2$.

Theorem 7.4. *Let f be a holomorphic selfmap of $\mathbf{P}^2$ of degree $d\geq 2$ which preserves a family of lines passing through a given point O in $\mathbf{P}^2$. Then f has a unique measure of maximal entropy.*

Remark 7.5. Holomorphic mappings of the form of Theorem 7.4 have been studied earlier, in other contexts [FS2], [U, Section 3.3], [JW].

Proof. Let d be the algebraic degree of f. Then the topological degree of f is d^2 and by a result of Gromov [G] we have $h(f)=\log d^2$.

Let $X\simeq\mathbf{P}^1$ be the set of lines in $\mathbf{P}^2$ passing through O and let π be the natural projection $\mathbf{P}^2\setminus O\to X$. Then π semiconjugates f to a holomorphic mapping $g: X\circlearrowleft$ of (topological) degree d.

Let Y be $\mathbf{P}^2$ blown up at O and let $p: Y\to\mathbf{P}^2$ be the blow-up map. Note that the exceptional divisor $E=p^{-1}(O)$ can be identified with X. In fact, π extends to a holomorphic mapping, still denoted π, of Y onto X and the restriction of π to E is a biholomorphism of E onto X. Further, f can be lifted in a unique way to a holomorphic mapping $\tilde{f}: Y\circlearrowleft$, such that π semiconjugates $\tilde{f}$ to $g: X\circlearrowleft$.

Now Y is a ruled surface and $\tilde{f}$ is a holomorphic selfmap of Y, which fibers over $g: X\circlearrowleft$. The topological degrees of f and g are d^2 and d, respectively. By

Theorem 7.3, $\tilde{f}$ has a unique measure $\tilde{\mu}$ of maximal entropy $\log d^2$. Since $\tilde{\mu}$ is ergodic and E is completely invariant for $\tilde{f}$ we must have $\tilde{\mu}(E)=0$, because otherwise

$$h_{\tilde{\mu}}(\tilde{f}) \leq h(\tilde{f}, E) = h(g, X) = \log d < \log d^2 = h(\tilde{f}).$$

Let μ be the invariant measure for f defined by $\mu = p_* \tilde{\mu}$. Recall that p is a biholomorphism outside E. Since $\tilde{\mu}(E)=0$ it therefore follows that $h_\mu(f) = h_{\tilde{\mu}}(\tilde{f}) = \log d^2$. Thus μ is a measure of maximal entropy for f. Suppose that $\nu \neq \mu$ is another measure of maximal entropy for f. Then there exists an invariant probability measure $\tilde{\nu}$ for $\tilde{f}$ such that $p_* \tilde{\nu} = \nu$. We have $h_{\tilde{\nu}}(f) \geq h_\nu(f) = \log d^2$ so $\tilde{\nu} = \tilde{\mu}$ by the uniqueness of $\tilde{\mu}$. Thus $\nu = p_* \tilde{\nu} = p_* \tilde{\mu} = \mu$ and we are done. □

In particular, Theorem 7.4 covers the case of *polynomial skew products* on $\mathbf{C}^2$. Such mappings were studied by Heinemann [H1], [H2], [H3], and by the author [J].

Corollary 7.6. *Let f be a polynomial skew product on $\mathbf{C}^2$ of degree $d \geq 2$, i.e. $f(z,w) = (p(z), q(z,w))$, where p and q are polynomials of degree d, and q has nonzero w^d-term. Then f has a unique measure of maximal entropy.*

Proof. The extension of f to $\mathbf{P}^2$ is given by

$$f[z:w:t] = [t^d p(z/t) : t^d q(z/t, w/t) : t^d].$$

Thus f satisfies the assumptions in Theorem 7.4 with $O = [0:1:0]$. □

Acknowledgement. The author thanks Marius Dabija for help on ruled surfaces, and Dror Varolin for interesting discussions on the definition of fibered rational maps.

References

[AR] ABRAMOV, L. M. and ROKHLIN, V. A., The entropy of a skew product of measure-preserving transformations, *Vestnik Leningrad. Univ.* **17**:7 (1962), 5–13 (Russian). English transl.: *Amer. Math. Soc. Transl.* **48** (1966), 255–265.

[B] BOGENSCHÜTZ, T., Entropy, pressure, and a variational principle for random dynamical systems, *Random Comput. Dynam.* **1** (1992/93), 99–116.

[Bo] BOWEN, R., Entropy for group endomorphisms and homogeneous spaces, *Trans. Amer. Math. Soc.* **153** (1971), 401–414.

[Br] BROLIN, H., Invariant sets under iteration of rational functions, *Ark. Mat.* **6** (1965), 103–144.

[D] DABIJA, M., Böttcher divisors, *Preprint,* 1998.

[FS1] FORNÆSS, J. E. and SIBONY, N., Random iterations of rational functions, *Ergodic Theory Dynam. Systems* **11** (1991), 687–708.

[FS2] FORNÆSS, J. E. and SIBONY, N., Critically finite rational maps on $\mathbf{P}^2$, in *The Madison Symposium on Complex Analysis* (Nagel, A. and Stout, E. L., eds.), Contemp. Math. **137**, pp. 245–260, Amer. Math. Soc., Providence, R. I., 1992.

[FS3] FORNÆSS, J. E. and SIBONY, N., Complex dynamics in higher dimension, in *Complex Potential Theory* (Gauthier, P. M. and Sabidussi, G., eds.), pp. 131–186, Kluwer, Dordrecht, 1994.

[FS4] FORNÆSS, J. E. and SIBONY, N., Complex dynamics in higher dimension II, in *Modern Methods in Complex Analysis* (Bloom, T., Catlin, D., D'Angelo, J. P. and Siu, Y.-T., eds.), Ann. of Math. Stud. **137**, pp. 135–182, Princeton Univ. Press, Princeton, N. J., 1995.

[FW] FORNÆSS, J. E. and WEICKERT, B., Random iterations in $\mathbf{P}^k$, to appear in *Ergodic Theory Dynam. Systems.*

[FLM] FREIRE, A., LOPEZ, A. and MAÑÉ, R., An invariant measure for rational maps, *Bol. Soc. Brasil. Mat.* **14** (1983), 45–62.

[G] GROMOV, M., Entropy, homology and semialgebraic geometry, *Astérisque* **145–146** (1987), 225–240.

[H1] HEINEMANN, S.-M., Dynamische Aspekte holomorpher Abbildungen in $\mathbf{C}^n$, *Ph. D. Thesis,* Göttingen, 1994.

[H2] HEINEMANN, S.-M., Julia sets for holomorphic endomorphisms of $\mathbf{C}^n$, *Ergodic Theory Dynam. Systems* **16** (1996), 1275–1296.

[H3] HEINEMANN, S.-M., Julia sets of skew products in $\mathbf{C}^2$, *Kyushu J. Math.* **52** (1998), 299–329.

[HP] HUBBARD, J. H. and PAPADOPOL, P., Superattractive fixed points in $\mathbf{C}^n$, *Indiana Univ. Math. J.* **43** (1994), 321–365.

[J] JONSSON, M., Dynamics of polynomial skew products on $\mathbf{C}^2$, *Math. Ann.* **314** (1999), 403–447.

[JW] JONSSON, M. and WEICKERT, B., A nonalgebraic attractor in $\mathbf{P}^2$, to appear in *Proc. Amer. Math. Soc.*

[K] KIFER, Y., *Ergodic Theory of Random Transformations,* Progress in Probability and Statististics **10**, Birkhäuser, Boston, Mass., 1986.

[KS] KOLYADA, S. and SNOHA, L., Topological entropy of nonautonomous dynamical systems, *Random Comput. Dynam.* **4** (1996), 205–233.

[LW] LEDRAPPIER, F. and WALTERS, P., A relativised variational principle for continuous transformations, *J. London Math. Soc.* **16** (1977), 568–576.

[L] LYUBICH, M., Entropy properties of rational endomorphisms of the Riemann sphere, *Ergodic Theory Dynam. Systems* **3** (1983), 351–385.

[M] MAÑÉ, R., On the uniqueness of the maximizing measure for rational maps, *Bol. Soc. Brasil. Mat.* **14** (1983), 27–43.

[MP] MISIUREWICZ, M. and PRZYTYCKI, F., Topological entropy and degree of smooth mappings, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **25** (1977), 573–574.

[R] ROKHLIN, V. A., Lectures on the theory of entropy of transformations with invariant measures, *Uspekhi Mat. Nauk* **22**:5(137) (1967), 3–56 (Russian). English transl.: *Russian Math. Surveys* **22** (1967) 1–52.

[S1] SESTER, O., Étude dynamique des polynômes fibrés, *Ph. D. thesis,* Université de Paris-Sud, 1997.

[S2] SESTER, O., Hyperbolicité des polynômes fibrés, *Bull. Soc. Math. France* **127** (1999), 393–428.

[Su] SUMI, H., Skew product maps related to finitely generated rational semigroups, *Preprint,* 1999.

[U] UEDA, T., Complex dynamical systems on projective spaces, *Preprint,* 1994.

[W] WALTERS, P., *An Introduction to Ergodic Theory,* Graduate Texts in Math. **79**, Springer-Verlag, New York–Berlin, 1982.

[Y] YOUNG, L. S., Ergodic theory of differentiable dynamical systems, in *Real and Complex Dynamical Systems* (Branner, B. and Hjorth, P., eds.), pp. 293–336, Kluwer, Dordrecht, 1995.

Received December 17, 1998
in revised form September 17, 1999

Mattias Jonsson
Department of Mathematics
University of Michigan
Ann Arbor, MI 48109-1109
U.S.A.
email: mattiasj@umich.edu