

Algebra & Number Theory

Volume 17
2023
No. 11

On self-correspondences on curves

Joël Bellaïche



On self-correspondences on curves

Joël Bellaïche

À la toute fin du processus de révision, le comité éditorial d'Algebra & Number Theory a eu l'immense tristesse d'apprendre le décès de Joël Bellaïche. Joël faisait des maths comme il respirait, avec une immense joie. Les mathématiques n'étaient pas pour lui séparées de la vie, elles étaient la vie même — la vie vraiment vivante.

At the very end of the reviewing process, the editorial board of Algebra & Number Theory learned with great sadness that Joël Bellaïche had passed away. Joël did math as he breathed, with immense joy. For him mathematics was not disconnected from life; it was life itself, and his way of living it to the fullest.

We study the algebraic dynamics of self-correspondences on a curve. A self-correspondence on a (proper and smooth) curve C over an algebraically closed field is the data of another curve D and two nonconstant separable morphisms π_1 and π_2 from D to C . A subset S of C is *complete* if $\pi_1^{-1}(S) = \pi_2^{-1}(S)$. We show that self-correspondences are divided into two classes: those that have only finitely many finite complete sets, and those for which C is a union of finite complete sets. The latter ones are called *finitary*, and happen only when $\deg \pi_1 = \deg \pi_2$ and have a trivial dynamics. For a nonfinitary self-correspondence in characteristic zero, we give a sharp bound for the number of étale finite complete sets.

Introduction	1867
1. Self-correspondences	1870
2. Finite complete sets	1877
3. The exceptional set of a nonfinitary self-correspondence	1882
4. The operator attached to a self-correspondence	1891
Acknowledgements	1898
References	1898

Introduction

Let k be a field, and let C be a smooth, proper and geometrically irreducible curve over k . By a *self-correspondence* on C (defined over k),¹ we mean the data of a smooth and proper scheme D over k , such that every connected component of D is a geometrically irreducible curve, and two k -morphisms

MSC2020: 14A10, 37E99.

Keywords: algebraic curve, self-correspondence, algebraic dynamics.

¹We adopt the definition of [Bullett and Penrose 2001] and [Krishnamoorthy 2018]. In part of the literature, a self-correspondence is defined instead as a divisor in the surface $C \times C$. The two notions are equivalent. To get a divisor of $C \times C$ using our definition, take $(\pi_1 \times \pi_2)(D)$, with multiplicities if several components of D have the same image in $C \times C$. To get from a divisor Δ of $C \times C$ a self-correspondence according to our definition, take the union of the normalization of each component of Δ repeated according to multiplicity. Our definition makes clearer the concepts of *étale* or *equiramified* complete sets, which are central in our study.

π_1 and π_2 from D to C , nonconstant and separable on every connected component of D . We denote by (D, π_1, π_2) , or often simply by D , that self-correspondence.

Fixing an algebraic closure \bar{k} of k , the intuitive way to think of a self-correspondence is as a multivalued map from $C(\bar{k})$ to itself defined by polynomial equations with coefficients in k , namely the map $x \mapsto \pi_2(\pi_1^{-1}(x))$. We call this multivalued map the *forward map* of the self-correspondence D .

Self-correspondences generalize endomorphisms: given an endomorphism f of a curve C , one can think of it as the self-correspondence $D_f := (C, \text{Id}_C, f)$. Like endomorphisms, two self-correspondences D and D' on C can be composed into a self-correspondence DD' (see Section 1.10) whose forward map is the composition of the (multivalued) forward maps of D and D' . Thus, like endomorphisms, a self-correspondence can be iterated. Better than endomorphisms, a self-correspondence $D = (D, \pi_1, \pi_2)$ always has a *transpose* (which plays in part the role of an inverse), denoted by tD and defined by ${}^tD = (D, \pi_2, \pi_1)$. The forward-map of tD , namely $x \mapsto \pi_1(\pi_2^{-1}(x))$ is called the *backward map* of D .

Let us introduce our fundamental terminology. A *forward-complete* (resp. *backward-complete*) set is a subset S of $C(\bar{k})$ that is stable by the forward (resp. backward) map (i.e., $\pi_1^{-1}(S) \subset \pi_2^{-1}(S)$, resp. $\pi_2^{-1}(S) \subset \pi_1^{-1}(S)$), and a *complete* set is a set which is both backward and forward-complete. An *irreducible complete set* is a minimal nonempty complete set.

The aim of this article is to answer elementary questions about finite complete sets for self-correspondences, such as when can there be infinitely many finite complete sets? This is a basic and fundamental question on the dynamics of self-correspondence. There is a relatively extensive literature on the subject of dynamics of self-correspondences on curves and even, less often, algebraic varieties. Most of this literature is concerned about correspondences over the complex numbers, see for instance [Fatou 1922; Bullett 1988; 1991; 1992; Bullett and Penrose 1994; 2001; Bullett and Lomonaco 2020; Bharali and Sridharan 2016; Pakovich 1996; 1995; 2008; Dinh 2002] (on \mathbb{P}^1), [Dinh 2005] (on \mathbb{P}^k), [Dinh and Sibony 2006] (on general varieties), [Dinh et al. 2020] (on general curves), but also over number fields; see [Autissier 2004; Ingram 2017; 2019], finite fields [Hallouin and Perret 2014] and general fields [Truong 2020]. To the best of our knowledge, the question we have in mind has not been solved, and not even asked, except in some very particular cases by Pakovich (see [Pakovich 1995] and Remark 3.3.7 below). A partial exception is the recent article of Krishnamoorthy [2018], essentially the second chapter of his PhD thesis at Columbia University. Though the focus of the paper is different, as Krishnamoorthy is concerned with general correspondences rather than self-correspondences, we borrow several ideas and concepts from him, and we gladly acknowledge our debt to his work.

Our first main result concerning finite complete sets is the following (see Theorem 2.2.1).

Theorem 1. *A self-correspondence (D, π_1, π_2) on a curve C over k has infinitely many finite complete sets if and only if there exists a nonconstant k -morphism*

$$f : C \rightarrow \mathbb{P}_k^1$$

such that

$$f \circ \pi_1 = f \circ \pi_2.$$

The method of proof of Theorem 1 is number-theoretic. Specifically, we use the famous theorem of Mordell, Weil and Néron asserting that the group of rational points of an abelian variety over a finitely generated field is a finitely generated abelian group. Theorem 1 seems difficult to prove by purely algebro-geometric methods or by complex-analytic ones in the case $k = \mathbb{C}$.

As a trivial consequence of Theorem 1, one sees that as soon as a self-correspondence has infinitely many finite complete sets, then in fact all its irreducible complete sets are finite, and moreover they have cardinality bounded by some integer M . A self-correspondence satisfying this property will be said *finitary*; its dynamical study is essentially trivial, in the sense that it reduces to dynamical questions over finite sets.

It is clear that only self-correspondences for which $\deg \pi_1 = \deg \pi_2$ (such a self-correspondence is said to be *balanced*) can be finitary. This explains why the notion of being finitary does not appear in the classical theory of complex dynamics, where one consider endomorphisms of \mathbb{P}^1 , which as correspondences have $(\deg \pi_1, \deg \pi_2) = (1, d)$, the case $d = 1$ being excluded as trivial. Even among balanced self-correspondences, the notion of finitary self-correspondence is very restrictive. They are in practice exceptions to all interesting general statements about the dynamics of self-correspondences. Nonfinitary self-correspondences are the natural domain of study of the dynamics of self-correspondences. For instance, an interesting result on the existence of a canonical invariant measure for a balanced self-correspondence on a curve over \mathbb{C} has recently been proved by Dinh, Kaufmann and Wu [Dinh et al. 2020] but only under a quite restrictive condition on the self-correspondence (see Remark 2.2.6). We believe that their main result (that the iterated pull-back of every smooth measure converges to the canonical measure) holds for all nonfinitary correspondences, and we plan to come back to this question in a subsequent work.

For a nonbalanced correspondence, there are no étale finite complete set, and finitely many nonétale ones. Moreover one can give an upper bound (in terms of the genera and degrees of the curves and maps involved) on the size of their union (see Proposition 2.1.1).

Balanced nonfinitary correspondences are much more subtle: they may have (finitely many) étale and nonétale finite complete sets; it is easy to give a bound on the number of finite nonétale complete sets, but we do not know how to bound their size. Our main objective is to bound the number of finite étale complete sets. With methods similar to those of the proof of Theorem 1, we are able to offer a bound (which happens to be optimal) only in some specific cases: when k is algebraic over a finite field (see Proposition 3.2.1); when k is arbitrary but $C = \mathbb{P}_k^1$ (see Proposition 3.2.3); and two other results when D is symmetric, that is $D \simeq {}^t D$ (see Propositions 3.2.5 and 3.3.3). But for more general results we need different, operator-theoretic, methods.

A self-correspondence D over C defines in a natural way a k -linear endomorphism T_D of the field of rational functions $k(C)$ of C , see Section 4.1. Whenever S is a forward-complete set, T_D stabilizes the subring B_S of $k(C)$ of functions whose all poles are in S . If moreover π_1 is étale on S , T_D stabilizes as well the natural filtration $(B_{S,n})_{n \geq 0}$, “by the order of the poles” of B_S .

The dynamical study of that action of T_D on the filtered ring B_S , when S is in particular étale complete, was the original motivation of this work. In fact, Hecke operators appearing in the theory of modular forms are of this type. Surprising results concerning the dynamics of the operators T_D for self-correspondences

over finite fields have been obtained in a recent work by Medvedovsky [2018], and, applied to Hecke operators, those results provide a new and elementary proof of certain deep modularity result of Gouvêa and Mazur [1998]. We plan to come back to these questions on a subsequent work. But in this paper, we content ourselves to use the operators T_D to obtain new informations on the dynamics of D .

We say that D is *linearly finitary* if there is a monic polynomial Q in $k[X]$ such that $Q(T_D) = 0$ as endomorphisms of $k(C)$. That D is linearly finitary means that the dynamics of T_D is trivial, in the sense that it is similar to the one of an operator on a finite-dimensional vector space. We prove the following:

Proposition 1. *A finitary self-correspondence is linearly finitary. The converse is true in characteristic zero.*

The direct sense is very easy. We prove the converse using graph-theoretic methods and Theorem 1 (see Proposition 4.4.7).

Our second main result is the following (see Theorem 4.5.3 which is slightly more precise).

Theorem 2. *If a self-correspondence D has three irreducible étale finite complete sets, then it is linearly finitary. In particular, in characteristic zero, a nonfinitary correspondence has at most two irreducible finite étale complete sets.*

We give a brief description of the idea of the proof, which uses only elementary methods: the theorem of Riemann and Roch and linear algebra. The first étale finite complete set S is used to define, as above, the natural filtration $(B_{S,n})_{n \geq 0}$, “by the order of the poles” of B_S , which is stabilized by T_D . Is this filtration split as a $k[T_D]$ -filtration? In general, this has no reason to be true. But with a second irreducible étale complete set S' , one can show that this filtration is “almost split”, namely that there exists for every n a T_D -stable subspace $V_{S,S',n}$ in $B_{S,n+1}$ such that $B_{S,n} + V_{S,S',n} = B_{S,n+1}$ and $\dim V_{S,S',n}$ bounded independently of n .² We prove this by defining $V_{S,S',n}$ as the space of functions in $B_{S,n+1}$ that vanishes on S' at a suitable order, and using Riemann–Roch. Now a third finite étale complete set S'' give a second quasisplitting $V_{S,S'',n}$ of the filtration $B_{S,n}$. Using Riemann–Roch, we prove that the two filtrations are orthogonal, in the sense that $V_{S,S',n} \cap V_{S,S'',n} = 0$ for n large enough. Then, a linear algebra argument shows that all eigenvalues of T_D appearing in $B_{S,n+1}$ (for n large enough) already appear in $B_{S,n}$ and Theorem 2 follows.

1. Self-correspondences

1.1. Curves. Let k be a field. By a *curve* over k we shall mean a nonempty proper and smooth scheme over $\text{Spec } k$ which is equidimensional of dimension 1 and geometrically connected.

If C is a curve over k , we denote by $k(C)$ the function field of C . If C and C' are two curves over k , a nonconstant morphism of k -schemes $\pi : C \rightarrow C'$ is finite and flat, hence surjective and thus defines a

²The filtration $(B_{S,n})$ of B_S would be T_D -split if in addition of $B_{S,n} + V_{S,S',n} = B_{S,n+1}$ we required $B_{S,n} \cap V_{S,S',n} = 0$, or equivalently $\dim V_{S,S',n} = \dim B_{S,n+1}/B_{S,n}$. One has $\dim B_{S,n+1}/B_{S,n} \leq |S|$ for every n (with equality for n large enough), so requiring that $\dim V_{S,S',n}$ is bounded independently of n is a qualitative version of that property. Hence the phrase “almost split”.

morphism of k -extensions $\pi^* : k(C') \rightarrow k(C)$, which makes $k(C)$ a finite extension of $k(C')$. Explicitly, if $f \in k(C)$ is seen as a morphism from C to \mathbb{P}^1 , $\pi^* f = f \circ \pi$. We say that π is *separable* if $k(C)$ is a separable extension of $k(C')$.

1.2. Self-correspondences. Given a curve C over k , a *self-correspondence* (D, π_1, π_2) on C is the data of a noetherian reduced k -scheme D whose connected components are curves D_i over k , with two morphisms π_1 and π_2 from D to C whose restrictions to each D_i are nonconstant and separable. Often, the morphisms π_1 and π_2 will be implicit and we shall simply denote by D the self-correspondence (D, π_1, π_2)

Self-correspondences on a fixed curve C over k naturally form a category: a morphism from (D, π_1, π_2) to (D', π'_1, π'_2) is a surjective k -morphism $h : D \rightarrow D'$ such that $\pi'_i \circ h = \pi_i$ for $i = 1, 2$. In particular we have a notion of *isomorphism* of self-correspondences.

Example 1.2.1. Let k be a field of characteristic p (a prime or zero), $P(x, y) \in k[x, y]$ be a polynomial in two variables x and y , which has at least one nonzero monomial of the form $x^a y^b$ with $p \nmid a$ and one of the form $x^{a'} y^{b'}$ with $p \nmid b'$. Let D be the normalization of the projective curve defined by the affine curve D_0 of equation $P(x, y) = 0$. Then the maps $(x, y) \mapsto x$, and $(x, y) \mapsto y$, are algebraic functions on D_0 , and they define rational functions on D , or equivalently maps $\pi_1, \pi_2 : D \rightarrow \mathbb{P}^1$. These maps are nonconstant and separable in view of our condition on P . Thus (D, π_1, π_2) is a self-correspondence over \mathbb{P}^1 .

Example 1.2.2 (the arithmetic-geometric mean; see [Bullett 1991]). If a and b are two positive real numbers with $a < b$, we define two recurrence sequences $a_0 = a, b_0 = b$, and for $n \geq 1, a_n = \sqrt{a_{n-1}b_{n-1}}, b_n = (a_{n-1} + b_{n-1})/2$. Then it is well-known (Gauss) and elementary that (a_n) is strictly increasing, (b_n) is strictly decreasing, and both converge to a same positive real number $M(a, b)$ called *the arithmetic geometric mean of a and b* . Since the formulas giving a_n and b_n are homogenous of degree 1, nothing is lost by fixing all the a_n equal to 1. That is, if we define $c_n = b_n/a_n$ for any n , we have that $c_0 = b/a$ and for $n \geq 1, c_n = (1 + c_{n-1})/2\sqrt{c_{n-1}}$, and from c_n we can retrieve a_n by $a_n = a\sqrt{c_0} \cdots \sqrt{c_{n-1}}$ and $b_n = c_n a_n$. The sequence c_n is strictly decreasing and converges to 1.

Inspired by these classical facts, one define the arithmetic-geometric correspondence (D_{agm}, π_1, π_2) on $C = \mathbb{P}^1_{\mathbb{C}}$, with $D_{agm} = \mathbb{P}^1_{\mathbb{C}}, \pi_1(z) = z^2$ and $\pi_2(z) = (1 + z^2)/2z$. Thus $\pi_2(\pi_1^{-1}(z))$ is the multiset $\{(1 + z)/2\sqrt{z}\}$ where \sqrt{z} is allowed to be either of the two determinations of the square root of z . Its dynamics, studied in [Bullett 1991], is related to works of Gauss and his successors on the arithmetic-geometric mean.

Note that (D_{agm}, π_1, π_2) is a special case of the correspondences of Example 1.2.1, namely it is equivalent to the self-correspondence over \mathbb{P}^1 attached to the equation $4xy^2 = (x + 1)^2$.

Example 1.2.3. Let C be the complete Igusa curves of level N (with $(N, p) = 1$) over \mathbb{F}_p . Recall ([Gross 1990, pages 460–462], where the curve is denoted by $I_1(N)$) that C is the smooth completion of the affine Igusa curve, which is defined as the moduli space for triples (E, α, β) , where E is an elliptic curve over a scheme of characteristic p, α an embedding $\mu_N \hookrightarrow E$ and β an embedding $\mu_p \hookrightarrow E$.

For l a prime number not dividing Np , we define the *Hecke correspondence* D_l , moduli space for quadruples (E, α, β, H) where (E, α, β) is as above and H is a subgroup scheme of E locally of order l . Define $\pi_1 : D_l \rightarrow C$ as just forgetting H , and π_2 as sending (E, α, β, H) to $(E/H, \alpha', \beta')$ where α' and β' are defined in the obvious manner. Then (D_l, π_1, π_2) is a self-correspondence, called the Hecke correspondence at l , on the Igusa curve C .

1.3. Bidegree. If D is a finite disjoint union of curves $D = \coprod_i D_i$, C a curve, and $\pi : D \rightarrow C$ a map nonconstant on every component D_i of D , then we define $\deg \pi$ as $\sum_i \deg \pi|_{D_i}$.

The *bidegree* of a self-correspondence (D, π_1, π_2) on a curve C is the ordered pair of integers $(\deg \pi_1, \deg \pi_2)$. It is often denoted (d_1, d_2) . A self-correspondence is *balanced* when $d_1 = d_2$. For example, the bidegree of the arithmetic-geometric self-correspondence of Example 1.2.1 is $(2, 2)$, and the bidegree of the Hecke correspondence D_l of Example 1.2.3 is $(l + 1, l + 1)$.

1.4. Transpose. If (D, π_1, π_2) is a self-correspondence on C , so is (D, π_2, π_1) , called the *transpose* of (D, π_1, π_2) . We shall denote this correspondence by tD . Its bidegree is (d_2, d_1) , if the degree of D is (d_1, d_2) .

1.5. Self-correspondence of morphism type. Let f be a nonconstant separable morphism from C to C . We denote by D_f the self-correspondence (C, Id_C, f) . A self-correspondence D over C is *of morphism type* if it is isomorphic to some D_f . Equivalently, D is of morphism type if and only if its bidegree has the form $(1, d)$. Thus we see that the transpose of a self-correspondence of the form D_f is of morphism type only when f is an isomorphism, and in this case ${}^tD_f \simeq D_{f^{-1}}$.

1.6. Minimal self-correspondences. A self-correspondence is *minimal* if the map $\pi_1 \times \pi_2 : D \rightarrow C^2$ is generically injective, that is there are only finitely many points of C^2 such that the fiber of $\pi_1 \times \pi_2$ has more than one element. Equivalently, D is minimal if $k(D)$ is generated, as a k -algebra, by its two subfields $\pi_1^*(k(C))$ and $\pi_2^*(k(C))$.

For any self-correspondence D on C , there is a unique pair (D_{\min}, h) where D_{\min} is a minimal self-correspondence over C and $h : D_{\min} \rightarrow D$ a morphism of self-correspondences on C : take D_{\min} the normalization of the image of D by $\pi_1 \times \pi_2$.

A minimal self-correspondence is rigid, i.e., has a trivial group of automorphism. Indeed, an automorphism σ of a self-correspondence D stabilizes every fibers of $\pi_1 \times \pi_2$. If D is minimal, all those fibers but finitely many are singletons, so σ fixes all points of D but finitely many, which implies that $\sigma = \text{Id}_D$.

1.7. Symmetric self-correspondences. A self-correspondence D is *symmetric* if ${}^tD \simeq D$. Obviously a symmetric self-correspondence is balanced. A self-correspondence is symmetric if and only if there exists an automorphism η of the k -scheme D such that $\pi_1 \circ \eta = \pi_2$. When D is minimal, this automorphism η is necessarily an involution, since η^2 is an automorphism of the self-correspondence D , which is rigid.

The Hecke correspondence D_l on the Igusa curve (see Example 1.2.3) is symmetric with $\eta(E, \alpha, \beta, H) = (E/H, \alpha', \beta', E[l]/H)$.

1.8. Terminology on directed graphs. By a *directed graph* we shall mean the data $\Gamma = (V, Z, s, t)$ of two sets V and Z and two maps $s, t : Z \rightarrow V$. Elements of V are called *vertices*, elements of Z are called *edges*, and for $z \in Z$, $s(z)$ is the *source* and $t(z)$ the *target* of z . In particular, *self-loops* (i.e., edges z such that $s(z) = t(z)$) and *repeated edges* (i.e., edges $z_1 \neq z_2$ such that $s(z_1) = s(z_2)$ and $t(z_1) = t(z_2)$) are allowed.

We use the usual terminology for a directed graph: a *forward-neighbor* (resp. *backward-neighbor*) of a vertex x is a vertex y such that there is an edge z with source x and target y (resp. with source y and target x). More generally, we say that y is a *k-forward-neighbor* of x if $y = x$ when $k = 0$, or if y is a forward-neighbor of a $(k-1)$ -forward-neighbor when $k \geq 1$.

A subset S of V is said to be *forward-complete* (resp. *backward-complete*) if it contains every forward-neighbors (resp. backward-neighbors) of its vertices.³ A subset S of V is *complete* if it is both backward-complete and forward-complete. A complete subset S is *irreducible* if it is nonempty and has no complete proper nonempty subset.⁴

It is clear that the irreducible complete subsets of a directed graph are the connected components of the nondirected graph it defines, and that the complete sets are the union of connected components. A union and intersection of complete sets is complete, as is the complement of any complete set. Every complete set is a disjoint union of irreducible complete sets.

If x, y are in V , a *directed path of length n* from x to y is a sequences of n edges $p = (z_1, \dots, z_n)$ such that $s(z_1) = x$, $t(z_n) = y$ and $t(z_i) = s(z_{i+1})$ for $i = 1, \dots, n - 1$. A directed path from x to x is called a *directed cycle*.

We shall denote by $\text{np}_{x,y,n}$ the number of directed paths from x to y of length n .

If k is a ring, and $\Gamma = (V, Z, s, t)$ is a directed graph, we define the *adjacency operator* of Γ , $A_\Gamma : \mathcal{C}(V, k) \rightarrow \mathcal{C}(V, k)$ on the k -module $\mathcal{C}(V, k)$ of maps from V to k , by the formula

$$(A_\Gamma f)(y) = \sum_{z \in Z, t(z)=y} f(s(z)).$$

The matrix of A_Γ in the canonical basis of $\mathcal{C}(V, k)$ is the *adjacency matrix* of Γ . By induction, we check that if $x, y \in V$ and $n \geq 1$ an integer, then

$$(A_\Gamma^n(\delta_x))(y) = \text{np}_{x,y,n} . \tag{1}$$

We shall sometimes consider functions $f : V \rightarrow k \cup \{\infty\}$ where ∞ is a symbol not in k . For such a function, we define $A_\Gamma f$ by the same formula as above, with the convention that the sum of the right hand-side is ∞ if exactly one of its term is ∞ , and is undefined if two or more of its terms are ∞ .

³Another widely used terminology for the same notion is *forward-invariant*.

⁴An irreducible complete set S is sometimes called a *grand orbit* of the self-correspondence. Indeed, such a set S can be obtained by starting with any of its point and applying the forward (multivalued) map $\pi_2 \circ \pi_1^{-1}$ or the backward map $\pi_1 \circ \pi_2^{-1}$ repeatedly, which may explain this terminology.

1.9. The directed graph attached to a self-correspondence. If (D, π_1, π_2) is a self-correspondence of C over k , and \bar{k} is a fixed algebraic closure of k , we define the oriented graph Γ_D attached to D as $(C(\bar{k}), D(\bar{k}), \pi_1, \pi_2)$. It is clear that for a subset S of $C(\bar{k})$, the notions of being forward-complete, backward-complete, or complete, as defined in the introduction (first page), and the same notions as defined for a subset of a directed graph defined in Section 1.8, coincide.

If $z \in D(\bar{k})$ is an edge, we write $e_{i,z}$ for the index of ramification of π_i at z . An edge $z \in D(\bar{k})$ is said *ramification-increasing* (resp. *equiramified*, resp. *étale*) if $e_{1,z} \leq e_{2,z}$ (resp. $e_{1,z} = e_{2,z}$, resp. $e_{1,z} = e_{2,z} = 1$). We observe that there are only finitely many edges that are not étale (and a fortiori, not equiramified or ramification-increasing). This is because π_1 and π_2 are assumed separable.

A subset S of $C(\bar{k})$ is said *ramification-increasing*, *equiramified*, *étale* if all the edges whose both source and target are in S are *ramification-increasing*, etc.

For any vertex $x \in C(\bar{k})$, we have the formula

$$\sum_{z \in D(\bar{k}), \pi_1(z)=x} e_{1,z} = d_1, \tag{2}$$

$$\sum_{z \in D(\bar{k}), \pi_2(z)=x} e_{2,z} = d_2. \tag{3}$$

In particular, there are at most d_1 edges with source x and d_2 edges with target x , and the directed graph Γ_D is locally finite. Given a finite set of vertices $S \subset C(\bar{k})$, one has by summing the above formula

$$\sum_{z \in \pi_1^{-1}(S)} e_{1,z} = d_1 |S|, \tag{4}$$

$$\sum_{z \in \pi_2^{-1}(S)} e_{2,z} = d_2 |S|. \tag{5}$$

Remark 1.9.1. The directed graph of tD is the directed graph of D with source and target maps exchanged.

Lemma 1.9.2. *Let k be a finitely generated extension of a prime field. Then k has extensions of arbitrary large prime degrees.*

Proof. Let \mathbb{F} be the prime subfield of k , and let T_1, \dots, T_n be a transcendence basis of k over \mathbb{F} . Thus if $k_0 = \mathbb{F}(T_1, \dots, T_n)$, k has finite degree d over k_0 . The field k_0 admits extension of any prime degree p : if $n = 0$, then $k_0 = \mathbb{F}_p$ or \mathbb{Q} and the result is well-known, and if $n \geq 1$, for p prime, the polynomial $X^p - T_1$ has no root in k_0 hence is irreducible over k_0 ; see [Lang 2002, Theorem 9.1]. If $p \nmid d$, the composition of an extension of degree p of k_0 with k is an extension of k of degree p . □

Proposition 1.9.3. *The directed graph Γ_D has infinitely many irreducible complete sets. All but finitely many of them are étale.*

Proof. Let us consider C and D as embedded in a projective space over k (say \mathbb{P}_k^3), and let k_0 be the subfield of k generated over the prime subfield of by the coefficients of the projective equations of C and

D and the coefficients of the polynomials defining π_1 and π_2 . Replacing k by k_0 we may assume that k is of finite type over its prime subfield.

If k is a finite type extension of its prime subfield, and $x \in C(k)$ is a vertex, then if $z \in D(\bar{k})$ is an edge with source (resp. target) x , one has $z \in D(k')$ for k' some finite extension of k of degree $\leq d_1$ (resp. $\leq d_2$). Indeed, z belongs to the schematic fiber of π_1 at x , which is a finite k_1 -scheme of degree d_1 . It follows that any forward-neighbor (resp. backward-neighbor) of $x \in C(k)$ is defined on an extension of degree $\leq d_1$ (resp. $\leq d_2$) of k' . By induction, any vertex in the same irreducible complete set as x is defined over an extension of k of degree whose all prime factors are $\leq \max(d_1, d_2)$.

Given a finite family S_1, \dots, S_l of irreducible complete sets in $C(\bar{k})$, pick points $x_1 \in S_1, \dots, x_l \in S_l$. By replacing k by a finite extension, we may assume that x_1, \dots, x_l all belong to $C(k)$, and thus every points x in $S_1 \cup \dots \cup S_l$ belong to $C(k')$ for k' a finite extension of k (depending on x) of some degree whose all prime factors are less than $\max(d_1, d_2)$. By the above lemma, k has extensions of arbitrary large prime degrees, hence has an extension k'' of prime degree $p > \max(d_1, d_2)$ and C has a point whose field of definition contains k'' . Such a point cannot belong to $S_1 \cup \dots \cup S_l$, which shows that there are other irreducible complete sets in $C(\bar{k})$. Therefore the number of irreducible complete sets is infinite.

The second assertion is clear since there are only finitely many nonétale edges. □

Example 1.9.4. The directed graph of the Hecke correspondence D_l on the Igusa curve C (see Example 1.2.3) is well understood. Since D_l is symmetric, we loose no information by forgetting the orientation of the edges and looking at Γ_{D_l} as an undirected graph.

There are two obvious finite completes sets, the set of supersingular points and the sets of cusps. The complete set of supersingular points is étale (easy since $l \neq p$) and irreducible (this can be proved by direct analysis, or, as Krishnamoorthy notes in [Krishnamoorthy 2018], simply as a consequence of Proposition 3.2.5 below). The complete set of cusps may be reducible, and none of its irreducible components are étale (again a consequence of Proposition 3.2.5). The other complete sets are all infinite and have been called *isogeny volcanoes*: they consist in a cycle of some order n , with an infinite tree of valence $l + 1$ attached to each vertices of that cycle; see [Sutherland 2013; Kohel 1996].

1.10. Sum and composition of self-correspondences. Let (D, π_1, π_2) and (D', π'_1, π'_2) be two self-correspondences on a curve C , of bidegrees (d_1, d_2) and (d'_1, d'_2) .

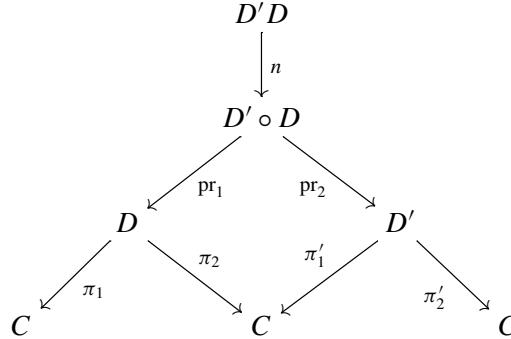
The *sum* of D and D' , denoted $D + D'$, is by definition the self-correspondence

$$\left(D \amalg D', \pi_1 \amalg \pi'_1, \pi_2 \amalg \pi'_2 \right)$$

on C . It is obvious that the oriented graph $\Gamma_{D+D'}$ has the same vertices as Γ_D and $\Gamma_{D'}$ and for set of edges the disjoint union of their set of edges. The bidegree of $D + D'$ is $(d_1 + d'_1, d_2 + d'_2)$.

We define $D' \circ D$ as the scheme $D \times_{\pi_2, C, \pi'_1} D'$. The two projections pr_1 and pr_2 of this fibered product over D and D' are finite and flat, and its total ring of fractions is étale over $k(C)$, since it is the tensor product of the two separable extensions $k(D)$ and $k(D')$ over $k(C)$ (they are seen as extensions of $k(C)$ through π_2^* and π'^*_1 respectively). In particular, $D' \circ D$ is proper over k , and all of its irreducible component

have dimension 1. We denote by $D'D$ the normalization of the reduced scheme attached to $D' \circ D$, and by $n : D'D \rightarrow D' \circ D$ the natural map:



Thus $D'D$ is a proper and smooth scheme of dimension 1, that is a disjoint union of curves. Since n is surjective, the restriction of $pr_1 \circ n$ and $pr_2 \circ n$ to every connected component of $D'D$ are surjective onto D and D' respectively and they are separable since n induces an isomorphism on the total rings of fractions. Thus, $(D'D, \pi_1 \circ pr_1 \circ n, \pi'_2 \circ pr_2 \circ n)$ is a self-correspondence on C , which we shall call *the composition of D' with D* . Its bidegree is $(d_1 d'_1, d_2 d'_2)$.

Example 1.10.1. If f, g are morphisms from C to C , then $D_f D_g = D_{f \circ g}$.

The directed graph $\Gamma_{D' \circ D}$ is the graph whose vertices are those of Γ_D or $\Gamma_{D'}$, and edges (z, z') where z is an edge of Γ_D , z' is an edge of $\Gamma_{D'}$ such that the source of z' is the target of z . The source (resp. target) of the edge (z, z') in $\Gamma_{D' \circ D}$ is the source of z in Γ_D (resp. the target of z' in $\Gamma_{D'}$). Hence

$$A_{\Gamma_{D' \circ D}} = A_{\Gamma_{D'}} \circ A_{\Gamma_D}. \tag{6}$$

Since n is surjective, generically an isomorphism, the directed graph $\Gamma_{D'D}$ is that of $\Gamma_{D' \circ D}$ described above, with finitely many new edges added, all those new edges having the same source and target that an already existing edge in $\Gamma_{D' \circ D}$.

Lemma 1.10.2. *If S is a forward-complete (resp. backward-complete, resp. complete) set for D and D' then it is forward-complete (resp. etc.) for $D'D$.*

If S is a complete étale set for D and D' , then the restrictions of the directed graphs of $D' \circ D$ and of $D'D$ to S coincide, S is also étale for $D'D$ and $A_{\Gamma_{D'D}}$ coincides with $A_{\Gamma_{D'}} \circ A_{\Gamma_D}$ on $\mathcal{C}(S, k)$.

Proof. The first assertion follows from what was said above. For the second, if $x \in S$, the fibers of D and D' at x are both étale, and so is their tensor product $D_x \times_k D'_x$, which is $(D' \circ D)_x$, so $D' \circ D$ is étale at points above x . But then it is smooth, and $n : D'D \rightarrow D' \circ D$ is thus an isomorphism on some neighborhood of the points above x , so $D'D$ is also étale over x , and the last assertion follows from (6). \square

It is clear that the composition of self-correspondences is associative up to obvious canonical identifications. We denote by D^n the composition of n copies of the self-correspondence D .

2. Finite complete sets

2.1. The unbalanced case.

Proposition 2.1.1. *Let (D, π_1, π_2) be a self-correspondence over C of bidegree (d_1, d_2) with $d_1 < d_2$. Denote by g_D, g_C the genera of D and C .⁵ Then any finite backward-complete set S satisfies*

$$|S| \leq 2 \frac{g_D - d_2 g_C + d_2 - 1}{d_2 - d_1}.$$

Moreover if such a set S is ramification-decreasing, then it is empty.

Proof. For the first assertion,

$$\begin{aligned} |S|d_2 &= \sum_{z \in \pi_2^{-1}(S)} e_{2,z} && \text{(by formula (5))} \\ &= |\pi_2^{-1}(S)| + \sum_{z \in \pi_2^{-1}(S)} (e_{2,z} - 1) \\ &\leq |\pi_2^{-1}(S)| + \sum_{z \in D(\bar{k})} (e_{2,z} - 1) \\ &\leq |\pi_2^{-1}(S)| + 2g_D - 2d_2g_C + 2d_2 - 2 && \text{(by Hurwitz's formula)} \\ &\leq |\pi_1^{-1}(S)| + 2g_D - 2d_2g_C + 2d_2 - 2 && \text{(since } S \text{ is backward-complete)} \\ &\leq |S|d_1 + 2g_D - 2d_2g_C + 2d_2 - 2, \end{aligned}$$

from which the bound of the statement immediately follows.

Now if S is a ramification-decreasing backward-complete finite set,

$$\begin{aligned} |S|d_2 &= \sum_{z \in \pi_2^{-1}(S)} e_{2,z} \\ &\leq \sum_{z \in \pi_2^{-1}(S)} e_{1,z} && \text{(since } S \text{ is ramification-decreasing)} \\ &\leq \sum_{z \in \pi_1^{-1}(S)} e_{1,z} && \text{(since } S \text{ is backward-complete)} \\ &= |S|d_1 && \text{(by formula (4))} \end{aligned}$$

which under our assumption $d_2 > d_1$ implies $S = \emptyset$. □

We record for later use a consequence of the proof:

Scholium 2.1.2. Let (D, π_1, π_2) be a self-correspondence with $d_1 \leq d_2$. If S is a ramification-decreasing backward-complete finite set for D , then S is complete and equiramified. Moreover, if a self-correspondence admits a complete equiramified nonempty finite set, then it is balanced.

⁵The genus g_D of a finite disjoint union of curves $D = \coprod_i D_i$ is defined here as the sum of the genera of D_i . With this definition, and that of degree given in Section 1.3, Hurwitz formula is still valid for a map $\pi : D \rightarrow C$.

Proof. The second chain of inequalities in the proof of the proposition is an equality by assumption $d_1 \leq d_2$, hence all intermediate inequalities must be equalities. The rest is clear. \square

Applying the proposition (resp. the scholium) to tD , we get a dual statement concerning forward-complete sets in the cases $d_1 > d_2$ (resp. $d_1 = d_2$) that we let the reader make explicit. Combining the proposition and its dual statement, we get:

Corollary 2.1.3. *Let (D, π_1, π_2) be a self-correspondence over C of bidegree (d_1, d_2) with $d_1 \neq d_2$ (that is, D is unbalanced). Let $d = \max(d_1, d_2)$. Then any finite complete set S is not equiramified, in particular is not étale, and satisfies $|S| \leq 2(g_D - dg_C + d - 1)/|d_2 - d_1|$.*

2.2. Finitary self-correspondences.

Theorem 2.2.1. *For a self-correspondence D on a curve C over a field k , the following are equivalent:*

- (i) *There exists a nonconstant k -morphism $h : C \rightarrow \mathbb{P}_k^1$ such that $h \circ \pi_1 = h \circ \pi_2$.*
- (ii) *There exists a nonconstant \bar{k} -morphism $h : C_{\bar{k}} \rightarrow \mathbb{P}_{\bar{k}}^1$ such that $h \circ \pi_1 = h \circ \pi_2$.*
- (iii) *There is an integer M such that every irreducible complete set has cardinality less or equal than M .*
- (iv) *All irreducible complete sets of D are finite.*
- (v) *All irreducible complete sets of D but possibly finitely many are finite.*
- (vi) *There are infinitely many finite complete sets.*

Proof. When D is unbalanced, (i) and (ii) are false by additivity of the degree, and (iii), (iv), (v) and (vi) are false by Proposition 2.1.1. The assertions (i) to (vi) are thus equivalent. We may therefore assume that D is balanced of degree d .

The implication (i) \Rightarrow (ii) is trivial. For (ii) implies (iii), we just note that the fibers $h^{-1}(t)$, $t \in \mathbb{P}^1(\bar{k})$ are complete since $\pi_1^{-1}(h^{-1}(t)) = (h \circ \pi_1)^{-1}(t) = (h \circ \pi_2)^{-1}(t) = \pi_2^{-1}(h^{-1}(t))$, and they all have size $\leq \deg h$. Therefore every irreducible complete set has cardinality less or equal than $\deg h$, which gives (iii).

That (iii) implies (iv) and (iv) implies (v) is trivial; and (v) implies (vi) by Proposition 1.9.3.

It is also not hard to prove that (ii) implies (i), as in [Krishnamoorthy 2018, Proposition 3.8]. In fact, assuming (ii), we know that there exists a finite extension k' of k and a map h' from $C_{k'}$ to $\mathbb{P}_{k'}^1$ defined over k' such that $h' \circ \pi_1 = h' \circ \pi_2$. Let V be the Weil's restriction of scalars of $\mathbb{P}_{k'}^1$ to k , and $\tilde{h} : C \rightarrow V$ the k -morphism corresponding to h' according to the universal property of Weil's restriction. The image C' of C by $\tilde{h} \circ \pi_1 = \tilde{h} \circ \pi_2$ in V , with its reduced scheme structure, is a closed subscheme of V defined over k and which has positive Krull dimension. There is therefore a rational function $u \in k(V)$ that induces a nonconstant map on C' . Thus, if $h := u \circ \tilde{h}$, h is a k -morphism $C \rightarrow \mathbb{P}_k^1$ such that $h \circ \pi_1 = h \circ \pi_2$ and (i) holds.

It only remains to prove that (vi) implies (ii).

To prove this, we first reduce to the case where k is of finite type over its prime field.

In any case, there is a subfield k_0 of k , of finite type over the prime subfield of k such that C , D , π_1 and π_2 are defined over k_0 . Assuming (vi), that is that $C(\bar{k})$ contains infinitely many finite irreducible

complete sets, we must show that (vi) holds for k_0 , that is $C(\bar{k}_0)$ also contains infinitely many finite irreducible complete sets. If all the irreducible finite complete sets in $C(\bar{k})$ are in $C(\bar{k}_0)$, we are obviously done. Otherwise, there is one irreducible finite complete set S in $C(\bar{k})$ which is not in $C(\bar{k}_0)$.

Let k' be any subfield of \bar{k} containing k_0 and of finite type over k_0 such that all points of S are defined over k' . There obviously exist such k' , and since S is not included into $C(\bar{k}_0)$, k' is transcendental over k_0 . Let $V = \text{Spec } A$ be an affine algebraic variety over k_0 whose field of fraction is k' and such that S spreads out to V , that is that all points of S are defined over A . Then V has positive Krull's dimension, so it has infinitely many \bar{k}_0 -points. Any \bar{k}_0 -point t in V gives by specialization of S a finite complete set S_t in $C(\bar{k}_0)$. It follows that $C(\bar{k}_0)$ contains at least one finite irreducible complete S_0 in $C(\bar{k}_0)$. Replacing k_0 by a finite extension, we may assume that $S_0 \subset C(k_0)$. We may now replace C by $C - S(k_0)$, D by $D - \pi_1^{-1}(S_0) = D - \pi_2^{-1}(S_0)$, and π_1 and π_2 by their suitable restriction to have a new self-correspondence (C, D, π_1, π_2) defined over k_0 , but with now C and D affine. Let us write $C = \text{Spec } R$, where R is a k -algebra

We now redo the same reasoning from the beginning, namely we prove that we may assume that there is a irreducible finite complete set S' in $C(\bar{k})$ which is not contained in $C(\bar{k}_0)$. One can write $S' = \{x_1, \dots, x_n\}$ where each x_i is a k_0 -morphism from R to \bar{k} . We now define A as the k_0 -subalgebra of \bar{k} generated by the images of the x_i 's in k' , $V = \text{Spec } A$, and k' the field of fraction of A . We can thus see the x_i as k_0 -morphisms from R to A . With these notation, for $t \in V(\bar{k}_0) = \text{Hom}_{k_0}(A, \bar{k}_0) = \text{Hom}_{k_0}(k', \bar{k}_0)$, the specialization S_t of S is by definition the subset $\{t \circ x_1, \dots, t \circ x_n\}$. Now we claim that for $t \neq t' \in V(\bar{k}_0)$, one has $S_t \neq S_{t'}$ as subsets of $C(\bar{k}_0)$; for if $S_t = S_{t'}$, then the subalgebra of \bar{k} generated by the $t \circ x_i$ ($i = 1, \dots, n$) on the one hand, by the $t' \circ x_i$ ($i = 1, \dots, n$) on the other hand, are equal, and thus $t = t'$ on the subalgebra generated by the image of x_i ($i = 1, \dots, n$), which is A , and this means $t = t'$. We thus obtain infinitely many distinct finite complete subset in $C(\bar{k}_0)$, namely the S_t where t moves in the infinite set $V(\bar{k}_0)$.

Replacing k by k_0 , we may henceforth assume that k is of finite type over its prime field.

We now continue the proof of (vi) \Rightarrow (ii) assuming k is finitely generated over the prime field. Let k^{sep} and k^{perf} be the separable and perfect closures of k in \bar{k} . Then $\bar{k} = k^{\text{sep}} \otimes_k k^{\text{perf}}$ and $\text{Gal}(k^{\text{sep}}/k) = \text{Gal}(\bar{k}/k^{\text{perf}})$. We denote this Galois group by G .

Let J be the Jacobian of C over k . Since k is of finite type over its prime field, $J(k)$ is a finitely generated abelian group by the theorem of Néron [1952]; see also [Lang and Néron 1959].

Moreover, one has

$$J(k^{\text{perf}}) \otimes_{\mathbb{Z}} \mathbb{Q} = J(k) \otimes_{\mathbb{Z}} \mathbb{Q}. \tag{7}$$

Indeed, there is nothing to prove in characteristic 0, and in characteristic $p > 0$, $J(k^{\text{perf}}) = \cup_n J(k^{1/p^n})$, so by induction it suffices to prove that $pJ(k^{1/p}) \subset J(k)$. But

$$pJ(k^{1/p}) = VFJ(k^{1/p}) \subset VJ^{(p)}(k) \subset J(k) \tag{8}$$

where $J^{(p)} = J \otimes_{k, x \mapsto x^p} k$, $F : J \rightarrow J^{(p)}$ and $V : J^{(p)} \rightarrow J$ are the relative Frobenius and Verschiebung k -morphisms, which completes the proof of (7). (We refer the reader to [Edixhoven et al. 2012, Section 5.2], for instance, for the definitions of Frobenius and Verschiebung, and to [Edixhoven et al. 2012, Proposition 5.19] for the fundamental relation $p = VF$ used above. Since the map F raises the coordinates of a point to the power p , it sends $J(k^{1/p})$ into $J^{(p)}(k)$, hence the middle inclusion in (8).)

Let r be the finite dimension of $J(k^{\text{perf}}) \otimes_{\mathbb{Z}} \mathbb{Q}$.

By assumption, there are infinitely many finite complete sets in $C(\bar{k})$, and therefore infinitely many étale finite complete sets. Each of them is a finite subset of $C(\bar{k})$, hence has a finite G -orbit. By grouping the irreducible complete sets by G -orbits, we see that there are still infinitely many disjoint étale finite complete sets invariant by G . Let us chose $r + 2$ of them, say S_0, \dots, S_{r+1} . To every S_i we attach the Weil divisor

$$\Delta_i = \sum_{x \in S_i} [x] \in \text{Div } C_{\bar{k}}$$

and let $\delta_i = |S_i|$ be its degree. The $r + 1$ divisors $\delta_0 \Delta_i - \delta_i \Delta_0$ have degree zero, hence they define points in $\text{Pic}^0(C_{\bar{k}}) = J(\bar{k})$, where J is the Jacobian of C , and those points are G -invariant, hence in $J(k^{\text{perf}})$. Therefore those points are \mathbb{Q} -linearly dependent, hence \mathbb{Z} -linearly dependent, which means that there are integers n_i , not all zero, $i = 0, \dots, r + 1$, and a nonconstant k^{perf} -map $h : C_{k^{\text{perf}}} \rightarrow \mathbb{P}_{k^{\text{perf}}}^1$ such that

$$\sum_{i=0}^{r+1} n_i \Delta_i = \text{div } h.$$

Since the S_i are étale, for $j = 1, 2$ one has

$$\pi_j^* \Delta_i = \sum_{z \in \pi_j^{-1}(S_i)} [z],$$

and since the S_i are complete, it follows that

$$\pi_1^* \Delta_i = \pi_2^* \Delta_i.$$

It follows that

$$\text{div}(h \circ \pi_1) = \pi_1^* \text{div } h = \pi_2^* \text{div } h = \text{div}(h \circ \pi_2),$$

hence that there exists $\lambda \in (k^{\text{perf}})^*$ such that

$$\lambda h \circ \pi_1 = h \circ \pi_2.$$

This implies that (reasoning as in the proof of (ii) implies (iii)) if S is any complete set, $h(S)$ is stable by multiplication by λ and λ^{-1} . By assumption, there exists a finite complete S such that $h(S)$ is not contained in $\{0, \infty\}$. This implies that λ is a root of unity, and if $\lambda^n = 1$, replacing h by h^n gives (ii). This completes the proof of the implication (vi) implies (ii), hence of the theorem. □

Remember from the introduction that when the equivalent assertions of the preceding theorem are satisfied, we say that D is *finitary*.

Example 2.2.2. The arithmetic-geometric mean self-correspondence introduced in Example 1.2.2 is not finitary. To see it, it suffices to notice that the irreducible complete set S containing a real number $c > 1$ also contains all elements $c_0 = c$, $c_n = (1 + c_{n-1})/2\sqrt{c_{n-1}}$ for any n , and that the sequence (c_n) , being strictly decreasing, take infinitely many value. Thus S is infinite, and D_{agm} is not finitary.

Remark 2.2.3. A self-correspondence D is finitary if and only if its transpose tD is finitary. This is seen trivially on any of the assertion (i) to (vi).

Also, if k' is any field containing k , a self-correspondence D is finitary if and only if its base change $D_{k'}$ is finitary. The implication D finitary $\Rightarrow D_{k'}$ finitary is clear on (i), and its converse is clear on (iii) or on (iv).

Remark 2.2.4. A correspondence of morphism type, say D_f , is finitary if and only if f is an automorphism of finite order. Indeed, if D_f is finitary, then $\deg f = 1$ by (i) and f is an automorphism, whose action on the generic fiber of h is a bijection of a finite set, so some power of f acts trivially on the generic fiber of h , hence on C , and f has finite order. The converse is trivial.

Remark 2.2.5. The method of using Jacobians in the proof of the theorem is inspired by Krishnamoorthy, [2018, Chapter 9], and our condition (a) is inspired by the condition he calls “having a core” which gives the title to his article.⁶ For Krishnamoorthy, a (general, not self-) correspondence $(D, \pi_1 : D \rightarrow C, \pi_2 : D \rightarrow C')$ has a core if there exist nonconstant maps $f : C \rightarrow \mathbb{P}^1$, $g : C' \rightarrow \mathbb{P}^1$ such that $f \circ \pi_1 = g \circ \pi_2$. For a self-correspondence, his notion is much weaker than our notion of being finitary, where we require $f = g$. Indeed, finitary implies balanced, while there are plenty of unbalanced self-correspondences that have a core, in particular all those of morphism type. Even for balanced self-correspondences, having a core does not imply being finitary, as the example of D_f when f is an automorphism of infinite order shows (or for less trivial examples, a suitable étale self-correspondence on a curve C of genus 1, for such a correspondence always have a core in the sense of Krishnamoorthy, but in general is not finitary). In the case of a symmetric self-correspondence D , however, one can show that D is finitary if and only if it has a core in the sense of Krishnamoorthy: see Lemma 3.2.4.

Remark 2.2.6. One finds in the literature yet another property akin to “having a core” or “being finitary”, namely what Dinh, Kaufmann et Wu calls “weakly modular” in [Dinh et al. 2020]. They work in the case $k = \mathbb{C}$ and they say that a balanced self-correspondence D over a curve C is *weakly modular* if there are two probability measures m_1 and m_2 over $C(\mathbb{C})$ such that $\pi_1^*m_1 = \pi_2^*m_2$. To “complete the square”, let us say that D is *weakly finitary* if there is one probability measure m on $C(\mathbb{C})$ such that $\pi_1^*m = \pi_2^*m$.

⁶Krishnamoorthy himself is inspired by Mochizuki, which introduces the same notion in the hyperbolic case, under the name “having an hyperbolic core”; see [Mochizuki 1998]. In the case $k = \mathbb{C}$, a closely related notion has also been considered by Bullett, Penrose [2001, Section 2.5], and their coauthors, under the name of *separable* (self-)correspondence. This notion is equivalent to “having a core” in the minimal case. Apparently, the two sets of authors (Bullett et al., Mochizuki and Krishnamoorthy) were unaware of each other’s works.

Thus one has a following square of implications for D a balanced self-correspondence over \mathbb{C} , none of which being an equivalence:

$$\begin{array}{ccc}
 D \text{ is finitary} & \implies & D \text{ has a core} \\
 \Downarrow & & \Downarrow \\
 D \text{ is weakly finitary} & \implies & D \text{ is weakly modular}
 \end{array}$$

Note that any self-correspondence which has a nonempty finite complete set S is weakly finitary, hence weakly modular, as taking m the normalized counting measure on S shows. Thus to be weakly modular is a weak condition.

Theorem 1.1 of [Dinh et al. 2020] states that for any balanced nonweakly modular self-correspondence D on a curve C over \mathbb{C} , there exists a measure μ_D on $C(\mathbb{C})$ which does not charge polar sets (in particular finite sets), and such that for every smooth measure μ on $C(\mathbb{C})$, $(\frac{1}{d}(\pi_1)_*(\pi_2^*)^n \mu) \rightarrow \mu_D$ as $n \rightarrow \infty$.⁷ In the case of a nonbalanced self-correspondence with $d_1 < d_2$, the theorem was known earlier [Bharali and Sridharan 2016]. One may ask whether this theorem holds more generally for any D that is not finitary. We conjecture the answer to be yes. (See also Remark 3.4.2.)

3. The exceptional set of a nonfinitary self-correspondence

3.1. Bounding the exceptional set. Let (D, π_1, π_2) be a self-correspondence over a curve C . The *exceptional set* E of C of D is the union of all finite complete sets of $C(\bar{k})$. Obviously, by Theorem 2.2.1, D is nonfinitary if and only if E is finite. But we may ask if in this case one can give an effective bound on the size of E . Unfortunately, the proof of Theorem 2.2.1 does not provide such a bound, even involving not only g_C, g_D , and g , but the Mordell–Weil rank r of the Jacobian variety of C over k_0 , because the step where we group together finite complete sets to get G -invariant complete sets is not effective.

We can ask two different questions on the exceptional set: we can ask for an upper bound on the number of elements of E , and for an upper bound on the number of irreducible components of E , that is the number of complete irreducible finite sets for D . The second question reduces to bounding the number of irreducible components of E that are étale, because the number of nonétale ones is less than the number of critical values of π_1 and π_2 , which in turns is bounded, using Hurwitz’s formula, by $2g_D - 2 - d(2g_C - 2)$.

In the case of an unbalanced self-correspondence, the two questions are solved by Proposition 2.1.1. We therefore limit ourselves to the balanced case.

3.2. The number of irreducible complete finite sets, I. Given a balanced nonfinitary self-correspondence D over C of bidegree (d, d) , can we give a bound to the number of étale irreducible complete finite sets

⁷We recall the standard measure-theoretic notation used here: let $\pi : D(\mathbb{C}) \rightarrow C(\mathbb{C})$ be an holomorphic map, nonconstant on every component of D ; if μ is a Borel measure on $D(\mathbb{C})$, then $\pi_*\mu$ is the measure on $C(\mathbb{C})$ defined by $\pi_*(\mu)(B) = \mu(\pi^{-1}(B))$; if μ is a Borel measure on $C(\mathbb{C})$, then $\pi^*\mu$ is the Borel measure on $D(\mathbb{C})$ defined by $\int f d\pi^*\mu = \int \pi_* f d\mu$ for every continuous function f on $D(\mathbb{C})$, where $\pi_* f(x) = \sum_{z \in \pi^{-1}(x)} f(z)$.

in terms of the genera g_C and g_D of the curve involved and the degree d ? We shall give several such bounds in particular but important situations (namely the case where k is a finite field, or when $C = \mathbb{P}^1$, or when the correspondence D is symmetric) using effective variants of the proof of Theorem 2.2.1. In the next section, using completely different methods we give a general result in characteristic zero (and also under a weaker form in characteristic p), namely that a nonfinitary correspondence has at most 2 étale (or even equiramified) finite irreducible sets (see Section 4.5).

Proposition 3.2.1. *Assume that k is algebraic over a finite field. If a correspondence D on a curve C over k is not finitary, then it has at most one nonempty finite equiramified complete set, and in particular at most one étale nonempty complete set.*

Note that the proposition implies that in case there is one nonempty equiramified complete set, it is automatically irreducible. Of course, D can very well have no nonempty finite complete set, as in the case of D_f , where f is a the translation by a nontorsion element on an elliptic curve.

Remark 3.2.2. The hypothesis made on k is necessary. Indeed, assume that k is not algebraic over a finite field. Then k has an element $t \neq 0$ which is of infinite multiplicative order (take an element which is transcendental over \mathbb{F}_p if k has characteristic p and $t = 2$ if k has characteristic 0). Consider the map $f(x) = tx$ from \mathbb{P}^1 to \mathbb{P}^1 , and the correspondence D_f it defines. This correspondence is of bidegree $(1, 1)$, and has two complete finite sets, obviously étale: $\{0\}$ and $\{\infty\}$. Yet D_f is not finitary, because f is not an automorphism of finite order.

Proof. Suppose that there are two distinct finite equiramified nonempty complete sets S and S' . If $S \subset S'$, replace S' by $S' - S$. If $S \not\subset S'$, replace S by $S - (S' \cap S)$. This way we may assume that S and S' are not only distinct, but disjoint.

In view of our hypothesis on k , we may assume that the self-correspondence is defined over a finite field k_0 , that S and S' are subsets of $C(k_0)$, and that $\pi_1^{-1}(S)$ and $\pi_2^{-1}(S')$ are subsets of $D(k_0)$.

Let $\Delta_S = \sum_{s \in S} [s]$ and $\Delta_{S'} = \sum_{s \in S'} [s]$ be the effective Weil's divisors attached to S and S' . The divisor $|S'| \Delta_S - |S| \Delta_{S'}$ has degree zero, hence is torsion in $\text{Pic}_{k_0}(C)$ (since it is an element of the finite group $\text{Pic}_{k_0}^0(C)$, the group of k_0 -rational points of the Jacobian of C). Therefore, there exist n and m such that $n \Delta_S - m \Delta_{S'}$ is a principal divisor; in other words, there exists a rational function h in $k(C)$ such that $\text{div } h = n \Delta_S - m \Delta_{S'}$.

We claim that there exists $\lambda \in k^*$ such that $h \circ \pi_1 = \lambda h \circ \pi_2$ as functions in $k(D)$, up to multiplication by a nonzero scalar. Indeed, for $i = 1, 2$, $\text{div } h \circ \pi_i = \pi_i^*(n \Delta_S - m \Delta_{S'}) = \sum_{t \in \pi_i^{-1}(S)} n e_{i,t} [t] - \sum_{t \in \pi_i^{-1}(S')} m e_{i,t'} [t']$ where $e_{i,t}$ is the ramification index of π_i at t . Using that S and S' are equiramified complete sets, we see that $\text{div } h \circ \pi_1 = \text{div } h \circ \pi_2$, hence the claim.

Now since λ belongs to a finite field, there is an n such that $\lambda^n = 1$. Replacing h by h^n , we see that D satisfies assertion (i) of Theorem 2.2.1. □

Proposition 3.2.3. *Let k be any field. If D a self-correspondence over \mathbb{P}_k^1 which is not finitary, there are at most two irreducible finite equiramified complete sets, and when there are two of such, there is no other irreducible finite complete set at all.*

Proof. Arguing as in the case of a finite base field k , but using the fact that $\text{Pic}_k^0 \mathbb{P}^1$ is trivial, we see that if S and S' are two disjoint irreducible equiramified sets, there is a function $f : \mathbb{P}^1 \rightarrow \mathbb{P}^1$ with divisor $|S'| \Delta_S - |S| \Delta_{S'}$. Such a function satisfies $f \circ \pi_1 = \lambda f \circ \pi_2$ for some $\lambda \in k^*$, and also by definition $f^{-1}(\infty) = S$, $f^{-1}(0) = S'$. Since D is not finitary, λ is not a root of unity, and as in the proof of Theorem 2.2.1, we see that for T complete, $f(T)$ is stable by multiplication by λ , and if T is also finite, this implies $f(T) \subset \{0, \infty\}$. Since $f^{-1}(0) = S$ and $f^{-1}(\infty) = S'$ are irreducible, they are the two only irreducible complete sets. \square

Finally, we give two more results in the same vein but for symmetric self-correspondences. They are based on earlier results in the literature and the following lemma.

Lemma 3.2.4. *If a self-correspondence has a core in the sense of Krishnamoorthy (see Remark 2.2.5, or [Krishnamoorthy 2018, Definition 3.5]) then tDD is finitary. If moreover D is symmetric then it is itself finitary.*

Proof. If (D, π_1, π_2) has a core, that is if there exists $f, g : C \rightarrow \mathbb{P}^1$ such that $f \circ \pi_1 = g \circ \pi_2$, then the forward map of D sends fibers of f to fibers of g , and the backward map of D , i.e., the forward map of tD , sends fibers of g to fibers of f . Thus tDD preserves the fibers of f , and is therefore finitary, which proves the first assertion.

If moreover D is symmetric, then D^2 is finitary and has infinitely many irreducible complete finite sets. But every irreducible complete set S of D breaks down in at most two irreducible complete sets of D^2 , namely the set of points of S which are connected to a given point x_0 of S by a path of even length, and its complement if nonempty. Therefore D must have infinitely many finite complete sets, and is therefore finitary. \square

Proposition 3.2.5 (Krishnamoorthy). *If D is a nonfinitary symmetric self-correspondence on a curve C over a field k , then D has at most one irreducible finite equiramified complete set. If k has characteristic zero and π_1, π_2 are étale, then D has no irreducible finite complete set.*

Proof. Krishnamoorthy proves that if D is a self-correspondence without a core, it has at most one finite irreducible finite étale complete set; see [Krishnamoorthy 2018, Theorem 9.6]. In fact his proofs work with “étale complete” replaced with “equiramified complete”, as in the proofs of Proposition 3.2.1 and Proposition 3.2.3. By the lemma above, this implies the proposition. \square

Corollary 3.2.6. *Let D be a nonfinitary symmetric self-correspondence on a curve C over a field k . Let S be an irreducible finite complete étale set for D . Then the undirected graph of S (obtained by forgetting the orientation of the edges) is not bipartite.*

Proof. If the undirected graph Γ_S attached to S was bipartite, then the graph $\Gamma_{S,2}$ with the same set of vertices S but whose edges are paths of degree 2 in S would be disconnected. But D^2 is also not finitary (by Lemma 3.2.4), and its graph on the set of vertices S is $\Gamma_{S,2}$ (by Lemma 1.10.2, using that S is étale), in contradiction with Proposition 3.2.5. \square

Remark 3.2.7. Let $S \in C(\bar{k})$ be a finite complete set for a self-correspondence D on C . We shall say that the set S is *consistently ramified* if for every undirected cycle with edges z_1, \dots, z_n , the rational number $\prod_{i=1}^n (e_{2,z_i}/e_{1,z_i})^{\epsilon_i}$ is 1, where the signs ϵ_i (for $i = 1, \dots, n$) are defined to be $+1$ if the edge z_i has a compatible orientation with z_{i+1} (i.e., the target of z_i is the source of z_{i+1} or the source of z_i is the target of z_{i+1}) or -1 otherwise (we use the convention that $z_{n+1} = z_1$). If S is equiramified, it is consistently ramified, since all factors in the above product are 1; but clearly the converse is false.

We claim that Propositions 3.2.1, 3.2.3, and 3.2.5 are still true with the phrase “finite equiramified complete sets” replaced with “finite consistently ramified complete sets”. Indeed, if S is consistently ramified, it is easy to see that one can attach to every $s \in S$ a positive integer n_s such that for every $z \in \pi_1^{-1}(S) = \pi_2^{-1}(S)$, one has $n_{\pi_1(z)}e_{1,z} = n_{\pi_2(z)}e_{2,z}$. In the proof of Proposition 3.2.1 (for example), it suffices to change the definition of Δ_S to be the divisor $\sum_{s \in S} n_s [s]$, to obtain a divisor with support S and such that $\pi_1^* \Delta_S = \pi_2^* \Delta_S$, and the rest of the proof may remain unchanged.

3.3. Bounding the exceptional set.

Proposition 3.3.1 (Krishnamoorthy). *If (D, π_1, π_2) is a nonfinitary symmetric self-correspondence on a curve C over a field k of characteristic 0, and if π_1 and π_2 are étale, then $E = \emptyset$.*

Proof. Krishnamoorthy proves that if π_1 and π_2 are étale and D has no core, and k has characteristic 0, then $E = \emptyset$; see [Krishnamoorthy 2018, Corollary 9.2]. The result follows using Lemma 3.2.4. \square

We shall now prove a generalization of the above result. A self-correspondence is *critically finite* (see [Bullett 1992]) if every ramification point in $C(\bar{k})$ of π_1 or π_2 belongs to a finite complete set. In other words, the union E_{crit} of the irreducible nonétale complete sets is finite.

Example 3.3.2. For the arithmetic geometric mean correspondence (see Example 1.2.2), it is easy to see that $E_{\text{crit}} = \{0, -1, 1, \infty\}$ and this correspondence is critically finite.

Proposition 3.3.3. *Let k be a field of characteristic zero. Let D be a critically finite self-correspondence on C over k having no core (in particular, symmetric and nonfinitary). Assume that the curve D is irreducible. Then D has no nonempty étale finite complete set. In other words, $E = E_{\text{crit}}$.*

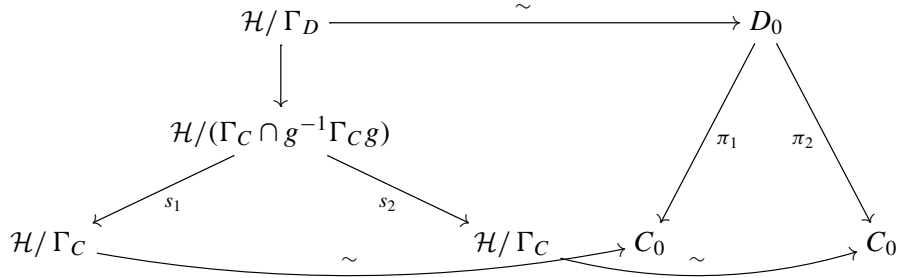
Proof. (Inspired by Section 3 of [Mochizuki 1998].)

We may assume that $k = \bar{k} = \mathbb{C}$. If E_{crit} is empty, then π_1, π_2 are étale, and the result follows from the Proposition 3.2.5. Let $C_0 := C - E_{\text{crit}}$, and $D_0 = D - \pi_1^{-1}(E_{\text{crit}}) = D - \pi_2^{-1}(E_{\text{crit}})$, and we still denote by π_1 and π_2 the restriction of π_1 and π_2 to D_0 . They are étale maps, and (D_0, π_1, π_2) is, in an obvious sense, a self-correspondence over C_0 in the category of open curves.

If $C = \mathbb{P}^1$, and E_{crit} has 1 or 2 elements, then C_0 is the affine line or punctured affine line, and it has at most one étale finite cover of any degree d . Thus $\pi_1 = \pi_2$ contradicting the assumption that D is not finitary.

In the remaining cases, the open curve $C_0 := C - E_{\text{crit}}$ is hyperbolic. Let $D_0 = D - \pi_1^{-1}(E_{\text{crit}}) = D - \pi_2^{-1}(E_{\text{crit}})$. Let us identify the universal cover of D_0 (which is also a universal cover of C_0) with the upper half-plane \mathcal{H} .

Fix some $x \in C_0$, z_1 in $\pi_1^{-1}(x) \in D_0$ and $h_1 \in \mathcal{H}$ a point that maps to z_1 in D_0 . The fundamental groups $\pi_1(C_0, x)$ and $\pi_1(D_0, z_1)$ are canonically identified, after those choices, with discrete subgroups of $\mathrm{PSL}_2(\mathbb{R})$ that we shall denote respectively by Γ_C and Γ_D ; we have $\Gamma_D \subset \Gamma_C$, the inclusion being of finite index. Choose also a $z_2 \in \pi_2^{-1}(z) \in D_0$, $h_2 \in \mathcal{H}$ that maps to z_2 , and a $g \in \mathrm{PSL}_2(\mathbb{R})$ such that $gz_1 = z_2$. Thus $\pi_1(D_0, z_2)$ is canonically identified with $g\Gamma_Dg^{-1}$, which also is a subgroup of finite index of Γ_C , and we have a commutative diagram



where the unnamed horizontal maps (curved or straight) are the identifications fixed above, and the diagonal maps s_1 and s_2 are given by respectively the inclusion and the conjugation by g of $\Gamma_C \cap g^{-1}\Gamma_Cg$ into Γ_C , the vertical map being given by the inclusion

$$\Gamma_D \subset \Gamma_C \cap g^{-1}\Gamma_Cg.$$

This inclusion shows that $\Gamma_C \cap g^{-1}\Gamma_Cg$ has finite index in Γ_C and thus that g belongs to the commensurator of Γ_C in $\mathrm{PSL}_2(\mathbb{R})$.

Let Γ be the closure of the subgroup of $\mathrm{PSL}_2(\mathbb{R})$ generated by Γ_C and $g^{-1}\Gamma_Cg$. We claim that Γ has infinite index in Γ_C . Otherwise, Γ would also be a lattice, and we would have two finite étale maps (surjective, of analytic stacks of dimension 1) $\mathcal{H}/\Gamma_C \rightarrow \mathcal{H}/\Gamma$ given by the inclusion and the conjugation by g^{-1} of Γ_C into Γ . We could then give an algebraic structure on \mathcal{H}/Γ , making it an algebraic stacks, and choose a nonconstant map of \mathcal{H}/Γ to \mathbb{P}^1 , which composed with the two finite étale maps $\mathcal{H}/\Gamma_C \rightarrow \mathcal{H}/\Gamma$ gives two morphisms of Riemann surfaces $f, g : \mathcal{H}/\Gamma_C = C_0 \rightarrow \mathbb{P}^1$ such that $f \circ \pi_1 = g \circ \pi_2$. Extending f, g to the complete smooth curves C , we thus see that the symmetric correspondence (D, π_1, π_2) has a core (f, g) , hence is finitary by Lemma 3.2.4, contradicting our hypothesis.

We just need to show that the self-correspondence of Riemann surfaces (D_0, π_0, π_1) has no finite complete sets, and by the commutativity of the diagram above it suffices clearly to show that the self-correspondence $\mathcal{H}/(\Gamma_C \cap g\Gamma_Cg^{-1}, s_1, s_2)$ on \mathcal{H}/Γ_C has no finite complete set. If S was such a complete set, its preimage \tilde{S} in \mathcal{H} would be invariant by Γ_D (obviously) and by $g\Gamma_Dg^{-1}$, hence by Γ . The set \tilde{S} would thus be an infinite union of Γ_C -orbits, contradicting the finiteness of S . □

We conclude this subsection by giving one result for self-correspondence on \mathbb{P}^1 in characteristic zero. It is a rephrasing of a result due to Pakovich.

Proposition 3.3.4. *Let k be a field of characteristic zero, (D, π_1, π_2) a self-correspondence over \mathbb{P}_k^1 of bidegree (d, d) . Let g_D be the genus of D . We assume that:*

- (i) *The singleton $\{\infty\}$ is a complete equiramified set.*
- (ii) *There exists a $\lambda \in k^*$ such that for every $z \in \pi_1^{-1}(\infty) = \pi_2^{-1}(\infty)$,*

$$\text{ord}_z(\pi_1 - \lambda\pi_2) > \text{ord}_z \pi_1 = \text{ord}_z \pi_2.$$

- (iii) *D is not finitary.*

Then $|E| \leq 3 + (2g_D - 1)/d$, with equality possible only in the case $g_D = 0, d = 1$.

Proof. This is essentially Théorème 1 of [Pakovich 1996]. More precisely, to prove our proposition we reduce to the case $k = \mathbb{C}$. Assuming that the proposition is false, there is a finite complete set in $\mathbb{P}_{\mathbb{C}}^1$ of cardinality $> 3 + (2g_D - 1)/d$, hence, removing ∞ , there is a finite complete set $K \subset \mathbb{C}$ of cardinality $> 2 + (2g_D - 1)/d$. Conditions (i) and (ii) allow us to apply Theorem 1 of [Pakovich 1996] which tells us that there is a rotation σ of the plane \mathbb{C} such that $\sigma(K) = K$ and $\pi_1 = \sigma \circ \pi_2$. Since K is finite, $\sigma^{(\#K)!}$ is the identity of K , hence of the real affine closure of K , and since $\#K \geq 2$, the rotation $\sigma^{(\#K)!}$ must fix a real line in \mathbb{C} , hence is the identity of \mathbb{C} . Thus $h(z) = z^{(\#K)!}$ satisfies $h \circ \pi_1 = h \circ \pi_2$, in contradiction with (iii). □

Remark 3.3.5. The condition (ii) is the most restrictive in practice. Assuming (i), it is clear that for every $z \in \pi_1^{-1}(\infty) = \pi_2^{-1}(\infty)$, there exists a $\lambda \in k^*$ such that $\text{ord}_z(\pi_1 - \lambda\pi_2) > \text{ord}_z \pi_1 = \text{ord}_z \pi_2$, but the existence of such a λ independent of $z \in \pi_1^{-1}(\infty)$ is problematic — except of course when $\pi_1^{-1}(\infty)$ is a singleton.

Moreover, the theorem is false without condition (ii), as the following example given by Pakovich shows: $D = \mathbb{P}^1$, $\pi_1(z) = (z^2 - z - 1)/(z^2 + z + 1)$, $\pi_2(z) = -(z^2 + 3z + 1)/(z^2 + z + 1)$. Then $\pi_1^{-1}(\{\infty\}) = \pi_2^{-1}(\{\infty\}) = \{j, j^2\}$ and $\{\infty\}$ is complete étale so (i) is satisfied, and one can check that (iii) is also satisfied. But $\{-1, 1\}$ is also a finite complete set (since $\pi_1^{-1}(\{-1, 1\}) = \{-1, 0, \infty\} = \pi_2^{-1}(\{-1, 1\})$). This provides an example where $C = D = \mathbb{P}^1$, and $|E| \geq 3$.

The following special case was proved by Pakovich [1995] in an earlier paper.

Corollary 3.3.6 (Pakovich). *Let k be a field of characteristic zero and suppose that $(\mathbb{P}_k^1, \pi_1, \pi_2)$ is a nonfinitary correspondence over \mathbb{P}_k^1 , where π_1 and π_2 are polynomials. Then $|E| \leq 2$.*

Proof. We can apply the proposition above: (i) is satisfied because $\pi^{-1}(\infty) = \pi^{-2}(\infty) = \{\infty\}$ and (ii) is satisfied because $\pi^{-1}(\infty)$ is a singleton. It tells us that $|E| \leq 3 + (2g_D - 1)/d = 3 - 1/d < 3$. □

Remark 3.3.7. This result can be considered as an effective version of Theorems 2.2.1 and 4.5.3 in a very special case; it states that for a nonfinitary correspondence $(\mathbb{P}^1, \pi_1, \pi_2)$ over \mathbb{P}^1 , with π_i polynomials, in characteristic zero, (thus having $S = \{\infty\}$ as equiramified complete set), then there is at most one other complete irreducible set S' , which moreover is also a singleton (if it exists). By contrast, our Theorem 4.5.3 below, under the same hypothesis will also tell us that there is at most one other complete

irreducible set, but without affirming that it has to be a singleton. However our conclusion will also hold without assuming that $C = D = \mathbb{P}^1$, nor any condition on π_1, π_2 and S .

3.4. Complements and questions.

Backward exceptional kernels. If $\Gamma = (V, E, s, t)$ is a directed graph, we define the *backward exceptional kernel* K_{backward} as the union in V of all finite backward-complete sets. A symmetric definition could of course be given for *forward exceptional sets* and we will let the interested reader reformulate the results below in this case.

If D is a self-correspondence over C of bidegree (d_1, d_2) , its *backward exceptional kernel* is the one of its associated directed graph Γ_D . A natural question for a self-correspondence is then: when is K_{backward} finite? We cannot offer a complete answer to this question. Here is what the author knows on K_{backward} .

First note that it is no restriction to assume that D is minimal, for there is always a minimal self-correspondence associated to D which has the same dynamics as D , in particular the same K_{backward} . So we assume below that D is minimal

When $d_1 < d_2$, Proposition 2.1.1 shows that K_{backward} is finite and gives a bound to its size.

What about K_{backward} when $d_1 \geq d_2$?

Consider first the case $d_2 = 1$, that is of a transpose of self-correspondence of morphism type: $D \simeq {}^t D_f$. If $d_1 = 1$, then f is an automorphism of C , of infinite order, and K_{backward} is finite. If $d_1 > 1$ however, then K_{backward} is always countable infinite. In the case where $d_1 > d_2 > 1$, K_{backward} may be finite or infinite: for instance, consider the case $D = C = \mathbb{P}^1$, $\pi_1(x) = x^3$, $\pi_2(x) = x^2$ of bidegree $(3, 2)$, where it is easy to see that $K_{\text{backward}} = \{0, \infty\}$, or for a balanced case, any symmetric nonfinitary correspondence (e.g., an Hecke correspondence D_l on the Igusa curve), where by symmetry, $K_{\text{backward}} = E$ which is finite. For an example with K_{backward} infinite consider the sum $D = {}^t D_f + {}^t D_g$, where f and g are the endomorphisms of \mathbb{P}^1 , $z \mapsto z^2$ and $z \mapsto z^3$. In this example K_{backward} is the set of all roots of unity while this self-correspondence is of bidegree $(5, 2)$, hence minimal since 5 and 2 are relatively prime.

Remains the balanced case $d_1 = d_2 > 1$. The finitary self-correspondences trivially have $K_{\text{backward}} = C(\bar{k})$. So let us consider a nonfinitary self-correspondence. Such a self-correspondence, if symmetric, will have $K_{\text{backward}} = E$ finite. The only question that remains is:

Question 3.4.1. Does there exist a nonfinitary balanced self-correspondence with K_{backward} infinite?

To analyze this question, note that for S a complete irreducible subset of $C(\bar{k})$, $S \cap K_{\text{backward}} = K_{\text{backward}}(S)$, and $K_{\text{backward}} = \coprod_S K_{\text{backward}}(S)$ when S run among irreducible complete subsets. If S is equiramified, then $K_{\text{backward}}(S)$ is a union of equiramified finite backward-complete subsets of S , and those subsets are complete by Scholium 2.1.2; thus, if $K_{\text{backward}}(S)$ is not empty it is finite and equal to S . This shows that the answer to the question is no when K_{backward} is equiramified, and because the number of nonequiramified irreducible complete sets S is finite, in general the question reduces to “is $K_{\text{backward}}(S)$ finite when S is a nonequiramified complete set?”

Backward exceptional and forward exceptional sets. If D is a self-correspondence over C of bidegree (d_1, d_2) , we define the *backward exceptional set* E_{backward} as the smallest forward-complete set containing K_{backward} . Be careful that the *backward exceptional set* is forward-complete by definition, but not in general backward-complete.

If D is not finitary, E_{backward} is always “small”: it contains the finite exceptional set E and is contained in the union of E and finitely many nonequiramified irreducible complete sets. In particular, E_{backward} is at most countable, and its complement contains an infinite union of irreducible complete sets. This follows easily from our study of K_{backward} .

If K_{backward} is étale, then it is complete and $E_{\text{backward}} = K_{\text{backward}}$ is finite. However, E_{backward} may be infinite in general. An example in bidegree $(2, 3)$ due to Dinh and Favre showing this is given in [Dinh 2005, Exemples 3.11]. A similar balanced example is given by $\pi_2(z) = z^2 - z, \pi_1(z) = z^2$, for which we have $K_{\text{backward}} = \{0, \infty\}$, but E_{backward} contains a strictly increasing sequence $0, 1, 2\sqrt{5}, \dots$ (each term x_{n+1} being the larger number in the pair $\pi_2(\pi_1^{-1}(x_n))$), hence is infinite.

Remark 3.4.2. The significance of the set E_{backward} appears most clearly in the ergodic theory of self-correspondences on curves over \mathbb{C} . More precisely, for a balanced nonweakly modular self-correspondence, it is shown in [Dinh et al. 2020, Theorem 1.2] that for every $x \in C(\mathbb{C}), (\frac{1}{d_2}(\pi_1)_*(\pi_2)^*)^n(\delta_x) \rightarrow \mu_D$, where the definition of the measure μ_D on C is recalled in Remark 2.2.6. Note that in the nonweakly modular case, it is easy to see that $K_{\text{backward}} = E_{\text{backward}} = \emptyset$. When $d_1 < d_2$, it is proved in [Dinh 2005; Dinh and Sibony 2006] that $(\frac{1}{d_2}(\pi_1)_*(\pi_2)^*)^n(\delta_x) \rightarrow \mu_D$ for every $x \notin E_{\text{backward}}$. It is therefore natural to conjecture that in every case $d_1 \leq d_2$, if D is not finitary, $(\frac{1}{d_2}(\pi_1)_*(\pi_2)^*)^n(\delta_x) \rightarrow \mu_D$ if and only if $x \notin E_{\text{backward}}$. At any rate, it is not hard to see that if $x \in E_{\text{backward}}$, if the limit $\lim_{n \rightarrow \infty} (\frac{1}{d_2}(\pi_1)_*(\pi_2)^*)^n(\delta_x)$ exists as a measure, then it charges at least one point in K_{backward} , and thus it cannot be μ_D .

Polarized self-correspondences.

Definition 3.4.3. Let (D, π_1, π_2) be a self-correspondence on a curve C over a field k . A *polarization* of D is an ample line bundle \mathcal{L} on C such that $\pi_1^*\mathcal{L}^n = \pi_2^*\mathcal{L}^m$ for some positive integers n and m . If D admits a polarization we say that D is *polarized*.

Recall that on a curve, a line bundle is ample if and only if its degree is positive. If \mathcal{L} is a polarization, one must have $d_1n = d_2m$.

A self-correspondence D has a polarization in each of the following cases:

- (i) $D = \mathbb{P}^1$ (in which case necessarily $C = \mathbb{P}^1$)
- (ii) k is algebraic over a finite field.
- (iii) D has a finite nonempty equiramified complete set $S \subset C(k)$.

Indeed, let us consider the group homomorphism $h : \text{Pic } C \rightarrow \text{Pic}^0 D, \mathcal{L} \rightarrow (\pi_1^*\mathcal{L})^{d_2} \otimes (\pi_2^*\mathcal{L})^{-d_1}$. Clearly D is polarizable if and only if $\ker h \not\subset \text{Pic}^0 C$. In particular, D is polarizable in case (i) since in this case $\text{Pic}^0 D = 0$ and $\ker h = \text{Pic } C = \mathbb{Z} \not\subset \text{Pic}^0 C = 0$). Also D is polarizable in case (ii) for in this case $\text{Pic}^0 D$

is torsion while $\text{Pic } C / \text{Pic}^0 C = \mathbb{Z}$ is torsion free. In case (iii), if the self-correspondence D has a finite equiramified complete set $\emptyset \neq S \subset C(k)$, take \mathcal{L} the line bundle attached to the divisor $\Delta_S = \sum_{s \in S} [s]$.

We now recall the basics of the general height theory of Lang and Néron, following the exposition of [Chambert-Loir 2011, Section 3] and [Serre 1989]. Assume that k is either a number field or a nontrivial finite type extension of an algebraically closed field. It is known that we can choose a family $M(k)$ of pairwise inequivalent absolute values on k and numbers $\lambda_v > 0$ such that for $a \in k^*$, $|a|_v = 1$ for almost all $v \in M(k)$ and the product formula holds: $\prod_{v \in M(k)} |a|_v^{\lambda_v} = 1$. Moreover, if k' is a finite extension of k , one can choose $M(k')$ in such a way that there is a surjective map $M(k') \rightarrow M(k)$ with finite fibers, such that for $a \in k$, $v \in M(k)$, $|a|_v^{\lambda_v} = \prod_{v' \in M(k'), \pi(v')=v} |a|_{v'}^{\lambda_{v'}}$, and the product formula holds. The choice of $M(k)$ allows one to define the *height function* on \mathbb{P}_k^n by

$$h([x_0, \dots, x_n]) = \log \left(\prod_{v \in M(k')} \max(|x_0|_v, \dots, |x_n|_v)^{\lambda_v} \right),$$

if $x_0, \dots, x_n \in k'$ (the result is independent of the choice of the finite extension k' containing x_0, \dots, x_n).

For a k -projective variety V , let us denote by $\mathcal{F}(V)$ the \mathbb{R} -vector space of maps from $V(\bar{k})$ to \mathbb{R} and by $\mathcal{F}b(V)$ the subspace consisting of these maps that are bounded. There is a unique morphism $\text{Pic } V \rightarrow \mathcal{F}(V) / \mathcal{F}b(V)$, $\mathcal{L} \mapsto h_{\mathcal{L}}$, such that if \mathcal{L} is very ample and ϕ is one of the embedding $V \rightarrow \mathbb{P}^n$ defined by \mathcal{L} , then $h_{\mathcal{L}} = h \circ \phi$. It follows that if $f : V \rightarrow W$ is a morphism of k -projective variety, $h_{f^*\mathcal{L}} = h_{\mathcal{L}} \circ f$.

Lemma 3.4.4. *If \mathcal{L} is ample, then $h_{\mathcal{L}} \neq 0$ in $\mathcal{F}(V) / \mathcal{F}b(V)$.*

Proof. We may assume that \mathcal{L} is very ample, and this reduces us to prove that for a projective subvariety in \mathbb{P}_k^n , the function h is unbounded on $C(\bar{k})$. Up to a linear change of variables, the map $\pi([x_0, \dots, x_n]) = [x_0, x_1]$ is surjective from V to \mathbb{P}^1 , and it is clear that $h(x) \geq h(\pi(x))$. It suffices therefore to show that h is unbounded on \mathbb{P}^1 , which is clear. □

Proposition 3.4.5. *Let (D, π_1, π_2) be an unbalanced polarized self-correspondence on a curve C over a field k that is not algebraic over a finite field. Then there are infinitely many vertices in Γ_D that do not belong to any directed cycle.*

Proof. Let \mathcal{L} be a polarization on D . The self-correspondence D together with \mathcal{L} are defined over a subfield k_0 of k which is of finite type over the prime subfield of k , and if k_1 is any field such that $k_0 \subset k_1 \subset \bar{k}$ it suffices obviously to prove the result for D considered as a self-correspondence over k_1 . Since \bar{k} is not the algebraic closure of a finite field, one can assume that k_1 is either a number field, or a nontrivial extension of finite type of an algebraically closed field. Replacing k_1 by k , we can now use the theory of height reminded above.

Let (d_1, d_2) be the bidegree of D . By symmetry of the statement to prove, we may and do assume that $d_1 < d_2$.

Since \mathcal{L} is a polarization of D , there exists an integer $n > 0$ such that $\pi_1^*(\mathcal{L}^{d_2 n}) = \pi_2^*(\mathcal{L}^{d_1 n})$. Thus $d_2 h_{\mathcal{L}} \circ \pi_1 = d_1 h_{\mathcal{L}} \circ \pi_2$ in $\mathcal{F}(D)/\mathcal{F}b(D)$. In other word, if h is any lift of $h_{\mathcal{L}}$ in $\mathcal{F}(C)$, there exists a positive real constant M such that for all $z \in D(\bar{k})$, $|d_2 h(\pi_1(z)) - d_1 h(\pi_2(z))| < M$.

If x_0, \dots, x_n is a directed cycle, then one has

$$\begin{aligned} |h(x_0) - (d_1/d_2)h(x_1)| &< M/d_2 \\ &\dots < \dots \\ |h(x_{n-1}) - (d_1/d_2)h(x_0)| &< M/d_2 \end{aligned}$$

so

$$|h(x_0) - (d_1/d_2)^n h(x_0)| < \frac{M}{d_2(1 + d_1/d_2 + \dots + (d_1/d_2)^{n-1})} < \frac{M}{d_2 - d_1}.$$

It follows that if x_0 belongs to a directed cycle then $h(x_0)$ is bounded by a constant (independent of the length of the cycle).

By the lemma, h is unbounded on $C(\bar{k})$. There is therefore infinitely many points in $C(\bar{k})$ that are not part of any directed cycle. □

Remark 3.4.6. The proposition is obviously false for a finitary self-correspondence but we do not know whether the proposition holds for a polarized balanced nonfinitary self-correspondence, nor whether the polarized hypothesis may be dropped (even in the unbalanced case).

4. The operator attached to a self-correspondence

4.1. Definition of the operator T_D . If C is a curve, and (D, π_1, π_2) a self-correspondence of C , we denote by T_{D, π_1, π_2} or simply T_D , the map $k(C) \rightarrow k(C)$ which sends f to

$$T_D f = \text{tr}_{k(D)/\pi_2^*k(C)} \pi_1^*(f).$$

The notations $\text{tr}_{k(D)/\pi_2^*k(C)}$ means the trace map from $k(D)$ to $k(C)$, where $k(D)$ is seen as an algebra over $k(C)$ of dimension d_2 through the map π_2^* . The map T_D is thus a k -linear endomorphism of $k(C)$.

4.2. Local description of the operator T_D . First recall some basic terminology. If C/k is a curve, or a disjoint union of curves, and if $f \in k(C)^*$, $x \in C(\bar{k})$, we denote by $\text{ord}_x(f)$ the order of vanishing of f at x . Thus $\text{ord}_x(f) > 0$ if $f(x) = 0$, $\text{ord}_x(f) = 0$ if $f(x) \in k^*$, and $\text{ord}_x(f) < 0$ if $f(x) = \infty$. By convention, we set $\text{ord}_x 0 = +\infty$. For S a subset of $C(\bar{k})$, we set

$$\text{ord}_S(f) := \inf_{x \in S} \text{ord}_x(f).$$

If $S = C(\bar{k})$, we simply write $\text{ord } f$ for $\text{ord}_S f$. One has $\text{ord } f \leq 0$ for any $f \in k(C)^*$, and $\text{ord } f = 0$ if and only if f is a nonzero constant.

Now suppose given a self-correspondence (D, π_1, π_2) of C over k .

Given $y \in C(\bar{k})$, and $z \in \pi_2^{-1}(y)$, we note K_y and K_z the fraction fields of the completions $\hat{\mathcal{O}}_{C,y}$ and $\hat{\mathcal{O}}_{D,z}$ of the local rings $\mathcal{O}_{C,y}$ and $\mathcal{O}_{D,z}$. The valuation ord_y on $k(C)$ (resp. ord_z on $k(D)$) extends uniquely

to K_y (resp. K_z), and $\mathcal{O}_{\widehat{C},y}$ (resp. $\mathcal{O}_{\widehat{D},z}$) is the ring of integers attached to this valuation; it is a complete discrete valuation ring. The map π_2^* induces an injective morphism $K_y \rightarrow K_z$, which makes K_z a finite separable extension of K_y , totally ramified of degree $e_{2,z}$ (which means that $\text{ord}_y(f) = e_{2,z} \text{ord}_z(f)$ for any $f \in K_y$ seen as an element of K_z).

Lemma 4.2.1. *Given $y \in C(\bar{k})$, $z \in \pi_2^{-1}(y)$ and $g \in K_z$, let $P_g(X) = X^{d_2} + \sum_{i=1}^{d_2} a_{i,g} X^{d_2-i} \in K_y[X]$ be the characteristic polynomial of the multiplication by g on K_z seen as a K_y -vector space of dimension d_2 through π_2^* . Then one has, for $i = 1, \dots, d_2 - 1$, $\text{ord}_z(a_{i,g}) \geq i \text{ord}_z(g)$ and $\text{ord}_z(a_{d_2,g}) = d_2 \text{ord}_z(g)$. In particular,*

$$\text{ord}_y \text{tr}_{K_z/K_y}(g) \geq \lceil \text{ord}_z(g)/e_{2,z} \rceil.$$

Proof. Since K_z/K_y is separable, $P_{g,z}(X) = \prod_{i=1}^{d_2} (X - \sigma_i(g))$ where the σ_i runs amongst the embedding on K_z into some normal closure L of K_z over K_y . If w is the valuation on L extending ord_z on K_z , then $w(\sigma_i(g)) = \text{ord}_z(g)$ and it follows that $\text{ord}_z(a_{i,g}) = w(a_{i,g}) \geq i w(g) = i \text{ord}_z(g)$, with equality if $i = d_2$. The last assertion follows since $a_{1,g} = \pm \text{tr}_{K_z/K_y}(g)$ and $\text{ord}_z(a_{1,g}) = e_{2,z} \text{ord}_y(a_{1,g})$. \square

Recall (Section 1.9) that $e_{1,z}$ and $e_{2,z}$ are the degrees of ramification of π_1 and π_2 at a point $z \in D(k)$.

Proposition 4.2.2. *Let $f \in k(C)$ and $y \in C(\bar{k})$. Set*

$$n = \min_{z \in D(k), \pi_2(z)=y} \left\lceil \frac{e_{1,z} \text{ord}_{\pi_1(z)} f}{e_{2,z}} \right\rceil.$$

Then

$$\text{ord}_y T_D f \geq n.$$

Moreover, if for any z such that $\pi_2(z) = y$, f has no pole at $\pi_1(z)$, then

$$f(y) = \sum_{z \in D(k), \pi_2(z)=y} e_{2,z} f(\pi_1(z)).$$

Proof. To compute the image of $T_D f$ in K_y we may extend the scalars from $k(C)$ to K_y since the formation of the trace commutes with base change. This means that $T_D f = \text{tr}_{K_y \otimes_{k(C)} k(D)/K_y} \pi_1^*(f)$ in K_y . But $K_y \otimes_{k(C)} k(D) = \prod_{z \in \pi_2^{-1}(y)} K_z$ (see, e.g., [Serre 1979, Chapter II, Section 3, Theorem 1]) hence

$$T_D f = \sum_{z \in \pi_2^{-1}(y)} \text{tr}_{K_z/K_y} \pi_1^*(f). \tag{9}$$

For $z \in \pi_2^{-1}(y)$, setting $x = \pi_1(z)$, we have $\text{ord}_z \pi_1^* f = e_{1,z} \text{ord}_x f$ hence by Lemma 4.2.1

$$\text{ord}_y \text{tr}_{K_z/K_y} \pi_1^*(f) \geq \left\lceil \frac{e_{1,z} \text{ord}_{\pi_1(z)} f}{e_{2,z}} \right\rceil \geq n.$$

By (9), $\text{ord}_y T_D f \geq n$.

To prove the second assertion, note that under its assumption, for any z such that $\pi_2(z) = y$, the image of $\pi_1^* f$ in K_z belongs to the complete d.v.r. $\mathcal{O}_{\widehat{D},z}$ and its image in the residue field \bar{k} is $f(\pi_1(z))$. Thus

$\text{tr}_{K_z/K_y} \pi_1^*(f) \in \hat{\mathcal{O}}_{D,z}$ and the image of that element in the residue field \bar{k} is $e_{2,z}f(\pi_1(z))$. The formula $f(y) = \sum_{z \in D(k), \pi_2(z)=y} e_{2,z}f(\pi_1(z))$ then follows from (9). \square

Recall (Section 1.8) that A_{Γ_D} is the adjacency operator of the graph D , and that for f a map from $C(\bar{k})$ to $\bar{k} \cup \{\infty\}$, $T_D(f)$ is a map from $C(\bar{k})$ minus a finite number of points to $\bar{k} \cup \{\infty\}$, the value of $T_D(f)$ being left undefined at any point where its computation requires to add the values of f at more than one pole.

Corollary 4.2.3. *Let $f \in k(C)$ seen as a map $f : C(\bar{k}) \rightarrow \bar{k} \cup \{\infty\}$. Then the functions $T_D f$ and $A_{\Gamma_D} f$ agree on all points of $C(\bar{k})$ but finitely many. They agree in particular on all étale complete sets on which f has no pole.*

Proof. Indeed, the last formula of the above proposition shows that the two functions agree at all points $y \in C(\bar{k})$ which are étale and not neighbors of a point where f has a pole. \square

Corollary 4.2.4. *if D' and D are two self-correspondence on C , $T_{D'D} = T_{D'} \circ T_D$.*

Proof. For $f \in k(C)$, $T_{D'D} f$ agrees almost everywhere with $A_{\Gamma_{D'D}} f$, which agrees almost everywhere with $A_{\Gamma_{D'}} A_{\Gamma_D} f$, which agrees almost everywhere with $T_{D'} T_D f$, and those two functions in $k(C)$ must then be equal. \square

4.3. The filtered ring B_S attached to a set of vertices S . Now fix a self-correspondence (D, π_1, π_2) on C over k and a set S of $C(\bar{k})$. We denote by $B_S \subset k(C)$ the rings of rational functions on C whose poles are all in S . Thus,

$$B_S = \{f \in k(C), \text{ord}_x(f) \geq 0 \text{ for all } x \in C(\bar{k}) - S\}.$$

If $S = \emptyset$, $B_S = k$. If S is not empty, the ring of fractions of B_S is $k(C)$. For $n \geq 0$, we set

$$B_{S,n} = \{f \in B_S, \text{ord}_S(f) \geq -n\}.$$

Lemma 4.3.1. *The k -subspaces $B_{S,n}$ for $n = 0, 1, \dots$ form an increasing exhaustive filtration of B_S . One has $B_{S,0} = k$. The quotients $B_{S,n}/B_{S,n-1}$ for $n \geq 1$ are spaces of dimensions $\leq |S|$, and of dimension exactly $|S|$ when S is finite and n is large enough relatively to $|S|$.*

Proof. The first two sentences are trivial. The last one follows from Riemann–Roch. \square

Proposition 4.3.2. *If S is forward-complete, the subring B_S of $k(C)$ is stable by T_D . If S is forward-complete and ramification-increasing, the filtration $(B_{S,n})$ is stable by T_D . The converses of both these statement hold if $\text{char } k = 0$ or $\text{char } k > d_2$.*

Proof. The first two statements follow from Proposition 4.2.2. The converse statements are left to the reader. \square

Remark 4.3.3. Remember (Proposition 2.1.1) that a forward-complete and ramification-increasing set S , nonempty and finite, may only exist if $d_1 \leq d_2$. In the balanced case $d_1 = d_2$, such a set S has to be complete and equiramified (Scholium 2.1.2).

Example 4.3.4. If C is the Igusa curve, and S the supersingular complete set, which is étale, B_S is the space of modular forms of level N , all weights, over $\bar{\mathbb{F}}_p$ and $(B_{S,n})$ is the *weight filtration* on that space; see [Gross 1990].

4.4. Linearly finitary self-correspondences. Let D be a self-correspondence on a curve C over k .

Definition 4.4.1. We say that D is *linearly finitary* if there is a nonzero polynomial $Q \in k[X]$ such that $Q(T_D) = 0$ on $k(C)$.

Proposition 4.4.2. *If D is finitary, then there is a monic polynomial $Q \in \mathbb{Z}[X]$ such that $Q(T_D) = 0$ on $k(C)$. In particular D is linearly finitary.*

Proof. If D is finitary, there exists an $M \geq 0$ such that every irreducible finite complete set of D is a directed graph with $\leq M$ vertices, with $\leq d_1$ (resp. $\leq d_2$) arrows starting (ending) at each point. There are finitely many such graphs up to isomorphism, so infinitely many irreducible complete sets S must be isomorphic to some finite directed graph Γ . If $Q(X)$ is the characteristic polynomial of the adjacency matrix of Γ , we see that for every $f \in k(C)$, $Q(T_D)f$ is zero on infinitely many irreducible complete sets, so $Q(T_D)f = 0$, and therefore $Q(T_D) = 0$. \square

Lemma 4.4.3. *Given two distinct points p, q in $C(\bar{k})$, a finite set $Z \subset C(\bar{k})$ not containing p or q , and an integer $n \geq 0$, there exists a rational function $f \in \bar{k}(C)$ such that $f(q) = 1$, f vanishes at every point of Z at order at least n , and f has no pole outside p .*

This follows from Riemann–Roch.

Lemma 4.4.4. *Let $Q(X) = \sum_{i=0}^n a_i X^i \in k[X]$ be a polynomial. The following are equivalent:*

- (i) *One has $Q(T_D) = 0$ in $k(C)$.*
- (ii) *There exists a nonempty forward-complete S such that $Q(T_D) = 0$ on B_S .*
- (iii) *There exists infinitely many irreducible étale complete sets S' such that $Q(A_{\Gamma_D}) = 0$ on $\mathcal{C}(S', k)$.*
- (iii') *There exists infinitely many irreducible étale complete set S' such that for every $x, x' \in S'$, one has $\sum_{i=0}^n a_i \text{np}_{x,x',i} = 0$ in k , where $\text{np}_{x,x',i}$ is the number of paths of length i from x to x' in Γ_D , as defined in Section 1.8.*
- (iv) *There exists an infinite étale complete set S' such that $Q(A_{\Gamma_D}) = 0$ on $\mathcal{C}(S', k)$.*
- (iv') *There exists an infinite étale complete set S' such that for every $x, x' \in S'$, one has $\sum_{i=0}^n a_i \text{np}_{x,x',i} = 0$ in k .*
- (v) *There exists an infinite backward-complete set S' such that $Q(A_{\Gamma_D})f$ has finite support for every $f \in \mathcal{C}(S', k)$.*

NB, the sets S in (ii) and S' in (iii) and (iii') are not assumed to be finite.

Proof. To prove that (i) implies (ii) is easy: take $S = C(\bar{k})$.

To prove that (ii) implies (iii), let p be a point in S . By Proposition 1.9.3, there exists infinitely many étale complete sets S' that do not contain p . Let S' be one of them, and let x be a point in S' . Let $\delta_x \in \mathcal{C}(S', k)$ be the function whose value at x is 1 and is 0 elsewhere. We provide Γ_D with the distance induced by its undirected graph structure. By Lemma 4.4.3, and since Γ_D is locally finite, there exists a function $f \in \bar{k}(C)$ such that $f = \delta_x$ on all points at distance $\leq 2n$ of x , and whose only possible pole is at p . In particular, $f \in B_S$, so $Q(T_D)f = 0$. By Corollary 4.2.3, $Q(A_{\Gamma_D})f = 0$ on S' . Clearly $Q(A_{\Gamma_D})f$ and $Q(A_{\Gamma_D})\delta_x$ agree on all points x' at distance $\leq n$ of x . Since $Q(A_{\Gamma_D})\delta_x$ has support in the sets of points at distance $\leq n$ of x , this implies that $Q(A_{\Gamma_D})\delta_x = 0$. Since this is true for an arbitrary point x of S' , $Q(A_{\Gamma_D}) = 0$ on $\mathcal{C}(S', \bar{k})$, hence (iii).

It is clear that (iii) implies (iv), by taking S' in (iv) to be the union of all the S' in (iii). Also, the equivalences between (iii) and (iii') and (iv) and (iv') is just Formula (1).

Finally, (iv) implies (v) is clear, so it just remains to prove that (v) implies (i). Let $f \in k(C)$. Since S' is backward-complete, $Q(A_{S'})f$ and $T_D f$ agree on every points of S' at distance $> n$ of a pole of f . Since the number of such points is finite, $Q(A_{S'})f$ and $T_D f$ agree on every point of S' except a finite number of them, and this implies that $T_D f$ has finite support on S' , and since S' is infinite, that $T_D f$ has infinitely many zeros. Hence $T_D f = 0$. □

Proposition 4.4.5. *If $Q(X) \in k[X]$ is a polynomial, then $Q(T_D) = 0$ if and only if $Q(T_D) = 0$. In particular, D is linearly finitary if and only if ${}^l D$ is. Moreover, if k' is any field extension of k , then D is linearly finitary if and only if $D_{k'}$ is linearly finitary.*

Proof. All is clear using condition (iii') or (iv') of Lemma 4.4.4. □

Lemma 4.4.6. *If S is an infinite irreducible étale complete set in Γ_D , then for every integer m and for every $x \in S$ there exists $x' \in S$ with a directed path from x to x' of length $m + 1$ and no directed path from x to x' of length $\leq m$.*

Proof. By symmetry of the statement to be proved we may assume that $d_1 \geq d_2$ to begin with. For $x \in S$, denote by F_x the smallest forward-complete set containing x , that is the set of all end points of directed paths starting at x . If F_x is finite, then it is complete by Scholium 2.1.2, contradicting the irreducibility of S which is infinite. Thus F_x is infinite.

Let $F_{x,m}$ be the subset of F_x consisting of all end points of directed paths of length $\leq m$ starting at x . Then $F_{x,m} \subset F_{x,m+1}$ and $F_x = \cup_m F_{x,m}$. If for some m , $F_{x,m} = F_{x,m+1}$, then $F_{x,m}$ is forward-complete, hence $F_x = F_{x,m}$ and F_x is finite, a contradiction. This for every m there exists $x' \in F_{x,m+1} - F_{x,m}$, which proves the lemma. □

Proposition 4.4.7. *If k has characteristic zero, and if D is linearly finitary, then D is finitary.*

Proof. Assume $Q(T_D) = 0$ for $Q \in k[X]$ a nonzero polynomial of degree n and dominant term $a_n \neq 0$. By Lemma 4.4.4, there exists infinitely many étale irreducible complete sets S' such that $Q(A_{S'}) = 0$. For S' any of them, we prove by contradiction that S' is finite. Indeed, choose $x \in S'$, and let x' be a point in S' with a directed path of length n from x to x' but no shorter path, which exists by the lemma above if

S' is infinite; then $0 = (Q(A_{S'})\delta_x) = a_n \text{np}_{n,x,x'}$, a contradiction since $\text{np}_{n,x,x'} \geq 1$ is nonzero in k . Thus D has infinitely many finite complete sets, and is therefore finitary. \square

Remark 4.4.8. This result is false in positive characteristic p . Take any self-correspondence D on a curve C which is nonfinitary. Then the self-correspondence $D' := D + D + \dots + D$ (p times) on C is still nonfinitary (it has the same complete sets as D), but the linear operator attached to this correspondence is $T_{D'} = pT_D = 0$, so D' is linearly finitary.

Corollary 4.4.9. *If D is linearly finitary, there is a monic polynomial $Q \in \mathbb{Z}[X]$ such that $Q(T_D) = 0$.*

Proof. There is a nonzero monic polynomial $Q(X) = \sum a_i X^i \in k[X]$ such that $Q(T_D) = 0$. Let k_0 be the prime subfield of k . Let l be a k_0 -linear form on k such that $l(1) = 1$ for some i . Then $Q_l(X) = \sum l(a_i)X^i \in k_0[X]$ is monic, and by the equivalence between (i) and (iii') in Lemma 4.4.4, one has $Q_l(T_D) = 0$. If k_0 is \mathbb{F}_p for some p , it suffices to lift Q_l into a monic polynomial in $\mathbb{Z}[X]$. If not, then k has characteristic zero, so D is finitary by Proposition 4.4.7, and there exists a monic polynomial in $\mathbb{Z}[X]$ that kills T_D by Proposition 4.4.2. \square

4.5. The number of irreducible complete finite sets, II.

Riemann–Roch calculations. Let k be an algebraic closed field, C a curve of genus $g \geq 1$ (just to avoid modifying the formulas in the case $g = 0$) and $S \subset C(k)$ a finite set. We denote as above by Δ_S the effective Weil divisor $\sum_{s \in S} [s]$ in $\text{Pic } C$.

Given a second finite nonempty set S' , disjoint from S , and any integer $n \geq 0$, there is clearly a unique relative integer $n' = n'(S, S', n)$ such that

$$2g - 2 + |S'| \geq (n - 1)|S| - n'|S'| > 2g - 2. \tag{10}$$

We denote by $V_{S,S',n}$ the subspace of $B_{S,n}$ of functions such that $\text{ord}_{S'} f \geq n'$ where n' is that integer, that is

$$V_{S,S',n} = \{f \in k(C), \text{div } f \geq -(n\Delta_S - n'\Delta_{S'})\}. \tag{11}$$

In plain English, $V_{S,S',n}$ is the space of all algebraic functions on C with poles of order at most n at every point of S , and at most n' at every point of S' . Remember that n' (defined by (10)) may be negative, so that the later requirement may mean that f has *zeros* of order at least $-n'$ on S' .

Lemma 4.5.1. *One has $B_{S,n-1} + V_{S,S',n} = B_{S,n}$ and $\dim V_{S,S',n} \leq g - 1 + |S| + |S'|$.*

Proof. The divisor $D = (n - 1)\Delta_S - n'\Delta_{S'}$ has degree $\deg D = (n - 1)|S| - n'|S'| > 2g - 2$ by assumption. Let $s_0 \in S$. By Riemann–Roch, $h(D + [s_0]) > h(D)$, so there exists a function $f_{s_0} \in k(C)$ such that $\text{div } f_{s_0} \geq -D - [s_0]$ and $\text{ord}_{s_0} f = -n$. Obviously $f_{s_0} \in V_{S,S',n}$ and the functions f_{s_0} , when s_0 runs in S , generate $B_{S,n}/B_{S,n-1}$. Hence the first assertion.

The second assertion also follows from Riemann–Roch, which says that

$$\begin{aligned} \dim V_{S,S',n} &= \deg(n\Delta_S - n'\Delta_{S'}) - g + 1 && (\text{since } \deg(n\Delta_S - n'\Delta_{S'}) > 2g - 2 \text{ and } g > 0) \\ &= |S| + (n - 1)|S| - n'|S'| - g + 1 \\ &\leq |S| + 2g - 2 + |S'| - g + 1 && (\text{by (10)}) \\ &g - 1 + |S| + |S'|. \end{aligned}$$

□

Now let S'' a third finite nonempty complete set, disjoint from S and from S' . Let n'' be defined by (10) with S' replaced by S'' , and let $V_{S,S'',n'}$ be defined similarly with (11). One has:

Lemma 4.5.2. *There exists an integer n_0 such that for $n > n_0$, $V_{S,S',n} \cap V_{S,S'',n} = 0$.*

Proof. The space $V_{S,S',n} \cap V_{S,S'',n}$ is the space of functions f such that $\text{div } f \geq -(n\Delta_S - n'\Delta_{S'} - n''\Delta_{S''})$. Using (10) for n' and n'' , one computes

$$\begin{aligned} \deg(n\Delta_S - n'\Delta_{S'} - n''\Delta_{S''}) &= n|S| - n'|S'| - n''|S''| \\ &= (n|S| - n'|S'|) + (n|S| - n''|S''|) - n|S| \\ &\leq 2g - 2 + |S'| + |S''| - n|S|. \end{aligned}$$

This number is negative for $n > (2g - 2 + |S'| + |S''|)/|S|$, and therefore $V_{S,S',n} \cap V_{S,S'',n} = 0$. □

The bound.

Theorem 4.5.3. *Let D be a self-correspondence on a curve C over an arbitrary field k . Assume that T_D is not linearly finitary. Then D has at most two irreducible complete equiramified finite sets.*

Proof. We may and do assume that k is algebraically closed. Also assume (just for simplicity in the formula) that the genus g of C is ≥ 1 , the case $C = \mathbb{P}^1$ being taken care of by Proposition 3.2.3. Assume that D admits three irreducible complete equiramified finite sets, S , S' and S'' . By Proposition 4.2.2, $V_{S,S',n}$ and $V_{S,S'',n}$ are stable by T_D . Let n_0 be as in Lemma 4.5.2. We claim that for $n > n_0$, every eigenvalue λ of T_D in $B_{S,n}$ is also an eigenvalue of T_D in $B_{S,n-1}$. We may assume that λ is an eigenvalue of T_D in the quotient $B_{S,n}/B_{S,n-1}$, otherwise there is nothing to prove. Let us call $m_\lambda \geq 1$ its multiplicity in $B_{S,n}/B_{S,n-1}$. By Lemma 4.5.1, λ is also an eigenvalue of T_D in V_{S,S'_n} (resp. in $V_{S,S'',n}$) with multiplicity $\geq m_\lambda$. Thus λ appears as an eigenvalue of T_D in $V_{S,S',n} + V_{S,S'',n}$ with multiplicity $\geq 2m_\lambda$, since the sum is direct by Lemma 4.5.2. Thus λ appears as an eigenvalue of T_D with multiplicity $\geq 2m_\lambda > m_\lambda$ in $B_{S,n}$, and it must appear in $B_{S,n-1}$.

By induction, all the eigenvalues of T_D on $B_{S,n}$ (hence on $V_{S,S',n}$) for any n already appear in B_{S,n_0} . Thus there are finitely many eigenvalues of T_D , say $\lambda_1, \dots, \lambda_l$, appearing in $V_{S,S',n}$ for any n . Define $Q(X) = (X - \lambda_1)^{g-1+|S|+|S'|} \dots (X - \lambda_l)^{g-1+|S|+|S'|}$. It is clear using Lemma 4.5.1 that $Q(T_D)$ kills $V_{S,S',n}$ for any n , hence, by Lemma 4.5.1, $B_{S,n}$ for any n , and thus $Q(T_D)$ kills B_S , and by Lemma 4.4.4, $Q(T_D)$ kills $k(C)$ contradicting the assumption that D is not linearly finitary. □

Corollary 4.5.4. *Let D be a self-correspondence on a curve C over a field k of characteristic zero. Assume that T_D is not finitary. Then D has at most two irreducible complete equiramified finite sets.*

This follows from the theorem and Proposition 4.4.7.

Remark 4.5.5. The upper bound, two, given by Theorem 4.5.3 for the number of irreducible complete equiramified finite sets for a nonlinearly finite self-correspondence is tight, in any characteristic: this is shown by the example considered in Remark 3.2.2, namely the self-correspondence attached to an automorphism of infinite order of \mathbb{P}_k^1 , when there exists one (which is the case if and only if k is not algebraic over a finite field).

Acknowledgements

The author would like to thank Antoine Chambert-Loir, Tien-Cuong Dinh, Anna Medvedovsky and two anonymous referees for useful comments.

References

- [Autissier 2004] P. Autissier, “Dynamique des correspondances algébriques et hauteurs”, *Int. Math. Res. Not.* **2004**:69 (2004), 3723–3739. MR Zbl
- [Bharali and Sridharan 2016] G. Bharali and S. Sridharan, “The dynamics of holomorphic correspondences of \mathbb{P}^1 : invariant measures and the normality set”, *Complex Var. Elliptic Equ.* **61**:12 (2016), 1587–1613. MR Zbl
- [Bullett 1988] S. Bullett, “Dynamics of quadratic correspondences”, *Nonlinearity* **1**:1 (1988), 27–50. MR Zbl
- [Bullett 1991] S. Bullett, “Dynamics of the arithmetic-geometric mean”, *Topology* **30**:2 (1991), 171–190. MR Zbl
- [Bullett 1992] S. Bullett, “Critically finite correspondences and subgroups of the modular group”, *Proc. Lond. Math. Soc.* (3) **65**:2 (1992), 423–448. MR Zbl
- [Bullett and Lomonaco 2020] S. Bullett and L. Lomonaco, “Mating quadratic maps with the modular group, II”, *Invent. Math.* **220**:1 (2020), 185–210. MR Zbl
- [Bullett and Penrose 1994] S. Bullett and C. Penrose, “Mating quadratic maps with the modular group”, *Invent. Math.* **115**:3 (1994), 483–511. MR Zbl
- [Bullett and Penrose 2001] S. Bullett and C. Penrose, “Regular and limit sets for holomorphic correspondences”, *Fund. Math.* **167**:2 (2001), 111–171. MR Zbl
- [Chambert-Loir 2011] A. Chambert-Loir, “Algebraic dynamics, function fields and descent”, lecture notes, 2011, available at http://www.math.u-psud.fr/~bouscare/workshop_diff/expose1.pdf.
- [Dinh 2002] T.-C. Dinh, “Ensembles d’unicité pour les polynômes”, *Ergodic Theory Dynam. Systems* **22**:1 (2002), 171–186. MR Zbl
- [Dinh 2005] T.-C. Dinh, “Distribution des préimages et des points périodiques d’une correspondance polynomiale”, *Bull. Soc. Math. France* **133**:3 (2005), 363–394. MR Zbl
- [Dinh and Sibony 2006] T.-C. Dinh and N. Sibony, “Distribution des valeurs de transformations méromorphes et applications”, *Comment. Math. Helv.* **81**:1 (2006), 221–258. MR Zbl
- [Dinh et al. 2020] T.-C. Dinh, L. Kaufmann, and H. Wu, “Dynamics of holomorphic correspondences on Riemann surfaces”, *Int. J. Math.* **31**:5 (2020), art. id. 2050036. MR Zbl
- [Edixhoven et al. 2012] B. Edixhoven, G. van der Geer, and B. Moonen, “Abelian varieties”, preprint, 2012, available at <http://van-der-geer.nl/~gerard/AV.pdf>.
- [Fatou 1922] P. Fatou, “Sur l’itération de certaines fonctions algébriques”, *Bull. Sci. Math.* **46** (1922), 188–198. Zbl
- [Gouvêa and Mazur 1998] F. Q. Gouvêa and B. Mazur, “On the density of modular representations”, pp. 127–142 in *Computational perspectives on number theory* (Chicago, 1995), edited by D. A. Buell and J. T. Teitelbaum, AMS/IP Stud. Adv. Math. **7**, Amer. Math. Soc., Providence, RI, 1998. MR Zbl

- [Gross 1990] B. H. Gross, “A tameness criterion for Galois representations associated to modular forms (mod p)”, *Duke Math. J.* **61**:2 (1990), 445–517. MR Zbl
- [Hallouin and Perret 2014] E. Hallouin and M. Perret, “Recursive towers of curves over finite fields using graph theory”, *Mosc. Math. J.* **14**:4 (2014), 773–806. MR Zbl
- [Ingram 2017] P. Ingram, “Critical dynamics of variable-separated affine correspondences”, *J. Lond. Math. Soc. (2)* **95**:3 (2017), 1011–1034. MR Zbl
- [Ingram 2019] P. Ingram, “Canonical heights for correspondences”, *Trans. Amer. Math. Soc.* **371**:2 (2019), 1003–1027. MR Zbl
- [Kohel 1996] D. R. Kohel, *Endomorphism rings of elliptic curves over finite fields*, Ph.D. thesis, University of California, Berkeley, 1996, available at <https://www.proquest.com/docview/304241260>.
- [Krishnamoorthy 2018] R. Krishnamoorthy, “Correspondences without a core”, *Algebra Number Theory* **12**:5 (2018), 1173–1214. MR Zbl
- [Lang 2002] S. Lang, *Algebra*, 3rd ed., Grad. Texts in Math. **211**, Springer, 2002. MR Zbl
- [Lang and Néron 1959] S. Lang and A. Néron, “Rational points of abelian varieties over function fields”, *Amer. J. Math.* **81** (1959), 95–118. MR Zbl
- [Medvedovsky 2018] A. Medvedovsky, “Nilpotence order growth of recursion operators in characteristic p ”, *Algebra Number Theory* **12**:3 (2018), 693–722. MR Zbl
- [Mochizuki 1998] S. Mochizuki, “Correspondences on hyperbolic curves”, *J. Pure Appl. Algebra* **131**:3 (1998), 227–244. MR Zbl
- [Néron 1952] A. Néron, “Problèmes arithmétiques et géométriques rattachés à la notion de rang d’une courbe algébrique dans un corps”, *Bull. Soc. Math. France* **80** (1952), 101–166. MR Zbl
- [Pakovich 1995] F. Pakovich, “Sur un problème d’unicité pour les polynômes”, preprint, 1995, available at <https://tinyurl.com/pakovichpoly>.
- [Pakovich 1996] F. Pakovitch, “Sur un problème d’unicité pour les fonctions méromorphes”, *C. R. Acad. Sci. Paris Sér. I Math.* **323**:7 (1996), 745–748. MR Zbl
- [Pakovich 2008] F. Pakovich, “On polynomials sharing preimages of compact sets, and related questions”, *Geom. Funct. Anal.* **18**:1 (2008), 163–183. MR Zbl
- [Serre 1979] J.-P. Serre, *Local fields*, Grad. Texts in Math. **67**, Springer, 1979. MR Zbl
- [Serre 1989] J.-P. Serre, *Lectures on the Mordell–Weil theorem*, Aspects Math. **E15**, Vieweg & Sohn, Braunschweig, 1989. MR Zbl
- [Sutherland 2013] A. V. Sutherland, “Isogeny volcanoes”, pp. 507–530 in *ANTS X: proceedings of the tenth Algorithmic Number Theory Symposium* (San Diego, CA, 2012), edited by E. W. Howe and K. S. Kedlaya, Open Book Ser. **1**, Math. Sci. Publ., Berkeley, CA, 2013. MR Zbl
- [Truong 2020] T. T. Truong, “Relative dynamical degrees of correspondences over a field of arbitrary characteristic”, *J. Reine Angew. Math.* **758** (2020), 139–182. MR Zbl

Communicated by Antoine Chambert-Loir

Received 2021-03-05 Revised 2022-04-22 Accepted 2022-08-18

Mathematics Department, Brandeis University, Waltham, MA, United States

Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR
Antoine Chambert-Loir
Université Paris-Diderot
France

EDITORIAL BOARD CHAIR
David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Jason P. Bell	University of Waterloo, Canada	Philippe Michel	École Polytechnique Fédérale de Lausanne
Bhargav Bhatt	University of Michigan, USA	Martin Olsson	University of California, Berkeley, USA
Frank Calegari	University of Chicago, USA	Irena Peeva	Cornell University, USA
J.-L. Colliot-Thélène	CNRS, Université Paris-Saclay, France	Jonathan Pila	University of Oxford, UK
Brian D. Conrad	Stanford University, USA	Anand Pillay	University of Notre Dame, USA
Samit Dasgupta	Duke University, USA	Bjorn Poonen	Massachusetts Institute of Technology, USA
Hélène Esnault	Freie Universität Berlin, Germany	Victor Reiner	University of Minnesota, USA
Gavril Farkas	Humboldt Universität zu Berlin, Germany	Peter Sarnak	Princeton University, USA
Sergey Fomin	University of Michigan, USA	Michael Singer	North Carolina State University, USA
Edward Frenkel	University of California, Berkeley, USA	Vasudevan Srinivas	Tata Inst. of Fund. Research, India
Wee Teck Gan	National University of Singapore	Shunsuke Takagi	University of Tokyo, Japan
Andrew Granville	Université de Montréal, Canada	Pham Huu Tiep	Rutgers University, USA
Ben J. Green	University of Oxford, UK	Ravi Vakil	Stanford University, USA
Christopher Hacon	University of Utah, USA	Akshay Venkatesh	Institute for Advanced Study, USA
Roger Heath-Brown	Oxford University, UK	Melanie Matchett Wood	Harvard University, USA
János Kollár	Princeton University, USA	Shou-Wu Zhang	Princeton University, USA
Michael J. Larsen	Indiana University Bloomington, USA		

PRODUCTION

production@msp.org
Silvio Levy, Scientific Editor

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2023 is US \$485/year for the electronic version, and \$705/year (+\$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

ANT peer review and production are managed by EditFLOW[®] from MSP.

PUBLISHED BY

 **mathematical sciences publishers**
nonprofit scientific publishing

<http://msp.org/>

© 2023 Mathematical Sciences Publishers

Algebra & Number Theory

Volume 17 No. 11 2023

On self-correspondences on curves JOËL BELLAÏCHE	1867
Fitting ideals of class groups for CM abelian extensions MAHIRO ATSUTA and TAKENORI KATAOKA	1901
The behavior of essential dimension under specialization, II ZINOVY REICHSTEIN and FEDERICO SCAVIA	1925
Differential operators, retracts, and toric face rings CHRISTINE BERKESCH, C-Y. JEAN CHAN, PATRICIA KLEIN, LAURA FELICIA MATUSEVICH, JANET PAGE and JANET VASSILEV	1959
Bézoutians and the \mathbb{A}^1 -degree THOMAS BRAZELTON, STEPHEN MCKEAN and SABRINA PAULI	1985
Axiomatizing the existential theory of $\mathbb{F}_q((t))$ SYLVY ANSCOMBE, PHILIP DITTMANN and ARNO FEHM	2013
The diagonal coinvariant ring of a complex reflection group STEPHEN GRIFFETH	2033