On deformation types of real elliptic surfaces

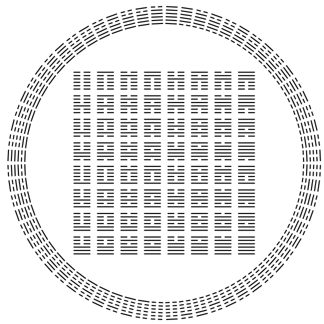
Alex Degtyarev, Ilia Itenberg, Viatcheslav Kharlamov

American Journal of Mathematics, Volume 130, Number 6, December 2008, pp. 1561-1627 (Article)

Published by Johns Hopkins University Press
DOI: *https://doi.org/10.1353/ajm.0.0029*

ON DEFORMATION TYPES OF REAL ELLIPTIC SURFACES

By Alex Degtyarev, Ilia Itenberg, and Viatcheslav Kharlamov



Le Yi Jing n'est pas un livre, un texte qu'on lit du début à la fin, mais un ouvrage que l'on consulte quand on en a besoin. Lorsqu'on hésite sur une voie à suivre, une attitude à prendre, un choix à faire, un dilemme à résoudre, on peut alors s'en servir pour ce qu'il est dans la pratique : un manuel d'aide à la décision.

Cyrille Javary, *Les Rouages du Yi Jing*, Ed. Phillipe Picquier, 2001

Abstract. We study real elliptic surfaces and trigonal curves (over a base of an arbitrary genus) and their equivariant deformations. We calculate the real Tate-Shafarevich group and reduce the deformation classification to the combinatorics of a real version of Grothendieck's *dessins d'enfants*. As a consequence, we obtain an explicit description of the deformation classes of M- and $(M-1)$- (i.e., maximal and submaximal in the sense of the Smith inequality) curves and surfaces.

Contents.

Manuscript received September 29, 2006.

Research of the second and third authors supported in part by the ANR-05-0053-01 grant of Agence Nationale de la Recherche and a grant of Université Louis Pasteur, Strasbourg.

American Journal of Mathematics 130 (2008), 1561–1627.

1. Introduction.

1.1. Motivation and historical remarks. In geometry of nonsingular algebraic surfaces, over the reals as well as over the complex numbers, there are two major equivalence relations: the first one, called *deformation equivalence*, is up to isomorphism and deformation (of the complex structure), and the second one, called *topological equivalence*, is up to diffeomorphism (ignoring the complex structure). Certainly, deformation equivalence implies topological equivalence, and one of the principal questions in the subject is to what extent the converse holds, i.e., to what extent is the deformation class of a surface controlled by its topology. Since we regard a real variety as a complex variety equipped with a real structure (which is an anti-holomorphic involution), by a deformation of real varieties we mean an equivariant Kodaira-Spencer deformation, and by a diffeomorphism between two real varieties we mean an equivariant diffeomorphism. Therefore, the Dif = Def question above stated over the reals would involve the same question for the underlying complex varieties. Luckily, due to Donaldson's and Seiberg-Witten's revolution in four dimensional topology (as well as the Enriques-Kodaira classification of algebraic surfaces), one does have an advanced level of control over the discrepancy between the deformation class of a compact complex surface and its diffeomorphism class. This fact makes it reasonable to fix a deformation class of compact complex surfaces beforehand and to concentrate on the topology and deformations of the real structures that can appear on (some of) the surfaces in question.

The problem of enumerating the equivariant deformation classes of real structures within a fixed complex deformation class goes back at least to F. Klein [Kl], who studied the nonsingular real cubic surfaces in $\mathbb{P}^3$ (i.e., Del Pezzo surfaces of degree 3) from a similar point of view. He proved that the equivariant deformation class of such a surface is already determined by the topology of its real part, which is the real projective plane with up to three handles or up to one sphere (Schläfli's famous five 'species' of nonsingular cubics). Further important steps in this direction were made by A. Commessatti [Co1], [Co2], who found a classification of all real abelian surfaces and all $\mathbb{R}$-minimal real rational surfaces, thus extending (at least implicitly) Klein's result to these special classes.

In general, we call a deformation class of complex varieties *quasi-simple* if a real variety within the complex class is determined up to equivariant deformation by the diffeomorphism type of the real structure. For curves, the problem was settled by F. Klein and G. Weichold (see, e.g., the survey [N1]) who proved that the family of compact curves of any given genus is indeed quasi-simple. Note that the equivariant deformation class of a real curve is no longer determined by its genus and real part; in addition, one should take into account the so called *type* of the curve, i.e., whether the real part does or does not divide the complexification. However, the type is certainly a topological invariant of the real structure.

Further advance in the study of quasi-simplicity called for appropriate tools in complex algebraic geometry. Their development took half a century, and it was not until the late 70s that the study was resumed. Now, due to the results obtained in [Ni], [DK2], [DIK1], [We], [CF], [DIK2], we know that quasi-simplicity holds for any special (in the sense of the Enriques-Kodaira classification) class of $\mathbb{C}$-minimal complex surfaces except elliptic. (For the surfaces of general type there are counter-examples, see, e.g., [KK].) A slightly different but related *finiteness* statement, i.e., finiteness of the number of equivariant deformation classes of real structures within a given deformation class of complex varieties, is known to hold for all surfaces except elliptic or ruled with irrational base. For ruled surfaces, the statement is probably true and its proof should not be difficult, cf., e.g., [DK2], but it does not seem to appear in the literature. Thus, elliptic surfaces are essentially the last special class of surfaces for which the quasi-simplicity and finiteness questions are still open.

It is worth mentioning that, in spite of noticeable activity in the theory of complex elliptic surfaces, literature dealing with the real case is scant. Among the few works that we know are [AMn], [Ba], [BMn], [DK1], [Fr], [GW], [Kh], [Mn], [Si], and [Wa].

1.2. Subject of the paper. In this paper, our goal is to study relatively minimal real elliptic surfaces without multiple fibers and, in particular, to understand the extent to which the equivariant deformation class of such a surface is controlled by the topology of its real structure. Recall that the complex deformation class of an elliptic surface as above is determined by the genus g of the base curve and the Euler characteristic χ of the surface, provided that χ is positive. (The case $g = 0$ is treated in A. Kas [Ka], and the general case, in W. Seiler [Se], see also [FM].) Note that, if χ is small (for a given genus), one deformation class may consist of several irreducible components: the principal component formed by the nonisotrivial surfaces may be accompanied by few others, formed by the isotrivial ones. Each isotrivial surface can be deformed to a surface that perturbs to a nonisotrivial one. However, from the known constructions it is not immediately obvious that the deformation can be chosen real. For this reason, we confine ourselves to the more topological study of nonisotrivial surfaces, leaving the algebro-geometric aspects to subsequent papers.

An elliptic surface comes equipped with an elliptic fibration. Moreover, for most surfaces, in particular, for all elliptic surfaces of Kodaira dimension 1, the elliptic fibration is unique. (In the case of relatively minimal surfaces without multiple fibers, the Kodaira dimension is known to be equal to 1 whenever $g > 0$, as well as when $g = 0$ and the Euler characteristic χ, which is divisible by 12, is > 24.) Thus, the elliptic fibration is an important part of the structure, and we include it into the setting of the problem, considering equivariant deformations of real elliptic fibrations (with no confluence of singular fibers allowed) on the one hand, and equivariant diffeo-/homeomorphisms on the other hand. Furthermore, as any nonisotrivial surface can be perturbed to an *almost generic* one, i.e., a surface with simplest singular fibers only, we consider solely deformations of almost generic surfaces. Here, "almost generic" can be thought of as "topologically generic", as opposed to "generic", or "algebraically generic", where one requires in addition that the fibers with nontrivial complex multiplication should also be simple. We use the latter assumption when treating an individual surface via algebro-geometric tools.

Note that during the deformation we never assume the base curve fixed; it is also subject to a deformation. The classification of real elliptic surfaces over a fixed base does not seem feasible; in general it may not even be possible to perturb a given surface to an almost generic one.

1.3. Tools and results. As in the complex case, the study of real elliptic surfaces is based upon two major tools: the real version of the Tate-Shafarevich group, which enumerates all real surfaces with a given Jacobian, and a real version of the techniques of *dessins d'enfants*, which reduces the deformation classification of nonisotrivial Jacobian elliptic surfaces (or, more generally, trigonal curves on ruled surfaces) to a combinatorial problem. We develop the two tools and, as a first application, obtain a rather explicit classification of the so called M- and

$(M-1)$-surfaces (i.e., those maximal and submaximal in the sense of the Smith inequality) and M- and $(M-1)$-curves. The principal results of the paper are stated in 6.2–6.4.

As a straightforward consequence of the description of deformation classes in terms of groups and graphs, the whole number of equivariant deformation classes of real elliptic fibrations with given numeric invariants is finite. This settles the finiteness problem stated above for nonisotrivial elliptic surfaces without multiple fibers.

The real Tate-Shafarevich group $\mathbb{R}Ш(J)$ is defined as the set (with a certain group operation) of the isomorphism classes of all real elliptic fibrations with a given Jacobian J. Contrary to the complex case, $\mathbb{R}Ш(J)$ is usually disconnected, and we describe, in purely topological terms, its discrete part $\mathbb{R}Ш^{\mathrm{top}}(J)$, which enumerates the deformation classes of real fibrations whose Jacobian is J. This description gives an explicit list of all modifications that a fibration may undergo, and the result shows that they can all be seen in the real part. As a consequence, we prove that, up to deformation, an elliptic surface is Jacobian if and only if the real part of the fibration admits a topological section (Proposition 4.3.5), each Betti number of an elliptic surface is bounded by the corresponding Betti number of its Jacobian, and each M- or $(M-1)$-surface is Jacobian up to deformation (Proposition 4.3.7).

The real version of *dessins d'enfants* was first introduced by S. Orevkov [Or1], who used it to study real trigonal curves on $\mathbb{C}$-minimal rational surfaces Σ_d. (A similar object was considered independently in [SV] and [NSV]). The curves considered by Orevkov do not intersect the 'exceptional' section of the surface; for even values of d, these are the branch curves of the Weierstraß models of Jacobian elliptic surfaces. (In the Weierstraß model, the elliptic surface appears as the double covering of Σ_d branched at the union of the exceptional section and a trigonal curve.) Using the dessin techniques, Orevkov invented a kind of Viro-LEGO® game: he introduced a few elementary pieces (which are the dessins of cubic curves), defined the operation of connecting "free ends" of two pieces, and used this procedure to construct bigger curves, thus proving a number of existence statements. Orevkov also noticed (private communication [Or2]; cf. similar observations in V. Zvonilov [Z]) that, as long as almost generic M-curves over a rational base are concerned, this procedure is universal, i.e., there is a unique way to break any M-curve into elementary pieces. Clearly, this construction gives a deformation classification of such M-curves.

We extend Orevkov's approach to trigonal curves over a base of an arbitrary genus and obtain similar results for M- and $(M-1)$-curves. We show that, as in the rational case, any M- or $(M-1)$- curve breaks into certain elementary pieces. (An essential ingredient here is Theorem 5.7.6, which states that unbreakable curves must be sufficiently "small". Another decomposability statement, Theorem 5.6.1, is used to handle large pieces of $(M-1)$-curves.) In the M-case, this procedure is unique; in the $(M-1)$-case it is unique up to a few moves that are described explicitly. As a consequence, we obtain a deformation classification of M- and $(M-1)$-curves and, when combined with the results on $\mathbb{R}Ш^{\mathrm{top}}$, that of M- and $(M-1)$-surfaces (see 6.3.1 and 6.3.2 for the M-case and 6.4.4 and 6.4.5 for the $(M-1)$-case). A surprising by-product of the classification is the fact that, essentially, M- and $(M-1)$-surfaces and curves exist only over a base of genus $g \leq 1$. (Here, "essentially" means that certain "trivial" handles should be ignored. Without this convention the genus can be made arbitrary large.)

1.4. Contents of the paper. Sections 2 and 3 are introductory. In Section 2 we remind the reader a few basic facts concerning topology of involutions, and in Section 3 we discuss certain complex and real aspects of the theory of trigonal curves, elliptic surfaces, their Jacobians and Weierstraß models. In Section 4 we introduce a real version of the Tate-Shafarevich group, express it in cohomological terms, and study its discrete part. The main results here are Theorems 4.2.7 and 4.3.2, as well as their corollaries. Section 5 plays a central rôle in the paper. Here we develop Orevkov's results on real *dessins d'enfants*. After a brief introduction, we concentrate on a special class of dessins that represent meromorphic functions having generic branching behavior, i.e., j-invariants of generic trigonal curves. The principal results of Section 5 are the decomposability Theorems 5.6.1 and 5.7.6, which assert that, under certain assumptions, a dessin breaks into simple pieces. Finally, in Section 6 we apply the results obtained to the case of M- and $(M-1)$-curves and surfaces. We prove the structure theorems, derive a few simple consequences, and discuss further generalizations and open problems.

1.5. *Acknowledgments.* Our thanks go to Stepan Orevkov, who enthusiastically shared his observations with us, motivating our interest in *dessins d'enfants*. We would also like to thank Victoria Degtyareva for courageously reading and polishing a preliminary version of the text.

We are grateful to the *Max-Planck-Institut für Mathematik* and to the *Mathematisches Forschungsinstitut Oberwolfach* and its RiP program for their hospitality and excellent working conditions which helped us to complete this project. An essential part of the work was done during the first author's visits to *Université Louis Pasteur*, Strasbourg.

2. Involutions and real structures. In this section we recall basic results concerning topology of involutions. Proofs and further details can be found in the monograph [Br1], which deals with general theory of compact transformation groups. A survey of sheaf theory, cohomology, and spectral sequences is found in [Br2]. For a self-contained exposition specially tailored for the needs of topology of real algebraic varieties, we refer to [DIK1].

2.1. Real structures and real sheaves. Throughout this section all topological spaces are assumed paracompact and Hausdorff.

2.1.1. A *real structure* on a complex variety X (not necessarily connected or nonsingular) is an anti-holomorphic involution $c_X\colon X \to X$. Clearly, any two real structures differ by an automorphism of X. By the Riemann extension theorem, an (anti-)holomorphic endomorphism f of the smooth part of X extends to an (anti-)holomorphic endomorphism of X if and only if f admits a continuous extension.

A *real variety* is a complex variety X equipped with a real structure c_X. (Sometimes it is convenient to refer to the pair (X, c_X) as a *real form* of X.) The fixed point set $\operatorname{Fix} c_X$ is called the *real part* of X and is denoted $X_{\mathbb{R}}$. A holomorphic map $f\colon X \to Y$ between two real varieties (X, c_X) and (Y, c_Y) is called *real* if it commutes with the real structures: $c_Y \circ f = f \circ c_X$.

Recall that for any continuous involution c_X on a finite dimensional topological space X with finitely generated total cohomology group $H^*(X;\mathbb{Z}_2)$ the following *Smith inequality* holds:

$$\dim H^*(\operatorname{Fix} c_X;\mathbb{Z}_2) \leq \dim H^*(X;\mathbb{Z}_2).$$

Furthermore, the difference $\dim H^*(X;\mathbb{Z}_2) - \dim H^*(\operatorname{Fix} c_X;\mathbb{Z}_2)$ is even. If the difference is $2d$, the involution c_X is called an $(M-d)$-*involution*. If c_X is the real structure of a real variety X, then X itself is called an $(M-d)$-*variety*.

2.1.2. Given an abelian group A with an involution $c\colon A \to A$, we define the cohomology groups $H^*(\mathbb{Z}_2;A)$ to be the cohomology of the complex

$$0 \longrightarrow A \xrightarrow{1-c} A \xrightarrow{1+c} A \xrightarrow{1-c} \dots$$

(the leftmost copy of A being of degree zero). Similarly, given a sheaf $\mathcal{A}$ with an involutive automorphism $c\colon \mathcal{A} \to \mathcal{A}$, we define the cohomology sheaves $\mathcal{H}^*(\mathbb{Z}_2;\mathcal{A})$ to be the cohomology of the complex

$$0 \longrightarrow \mathcal{A} \xrightarrow{1-c} \mathcal{A} \xrightarrow{1+c} \mathcal{A} \xrightarrow{1-c} \dots.$$

(Certainly, the former is nothing but a specialization of the general definition of the cohomology of a discrete group G with coefficients in a G-module, see, e.g., [Bro], to the group $G = \mathbb{Z}_2$ and the simplest invariant cell decomposition of the space $S^\infty = E\mathbb{Z}_2$. The latter is a straightforward sheaf version of the former.)

2.1.3. Let X be a topological space with an involution $c_X\colon X \to X$. Denote by $\pi\colon X \to X/c_X$ the projection. Given a sheaf $\mathcal{A}$ on X, any morphism $c\colon \mathcal{A} \to c_X^*\mathcal{A}$ (over the identity of X) descends to a morphism $\pi_*c\colon \pi_*\mathcal{A} \to \pi_*\mathcal{A}$. By a certain abuse of the language, c is called an *involution* if π_*c is an involution. This condition is equivalent to the requirement $c \circ c_X^*c = \mathrm{id}$, where $c_X^*c\colon c_X^*\mathcal{A} \to \mathcal{A}$ is the pull-back of c. (Certainly, this definition is merely an attempt to refer to involutive lifts $\mathcal{A} \to \mathcal{A}$ of c_X to $\mathcal{A}$ in terms of sheaf morphisms identical on the base.)

The constant sheaf G_X (for any abelian group G) has a canonical involution, which is the identity $G_X = c_X^* G_X$. As a lift of c_X it is given by $s \mapsto s \circ c_X$.

2.1.4. Now, let X be a complex manifold and let $c_X\colon X \to X$ be a real structure. Then the structure sheaf $\mathcal{O}_X$ has a canonical involution, called the *canonical real structure* (defined by c_X); it is given by the complex conjugation $\mathcal{O}_X = \overline{c_X^* \mathcal{O}_X}$, or, as a lift of c_X, by $s \mapsto \overline{s \circ c_X}$. If $\mathcal{A}$ is a coherent sheaf on X, the pull-back $c_X^* \mathcal{A}$ is a (coherent, in a sense) sheaf of $c_X^* \mathcal{O}_X$-modules. An involution $c\colon \mathcal{A} \to c_X^* \mathcal{A}$ is called a *real structure* on $\mathcal{A}$ if it is compatible with the module structure (via the canonical real structure on $\mathcal{O}_X$). A typical example is the canonical real structure on the sheaf $\mathcal{F}$ of sections of a *"Real"* vector bundle F (i.e., a holomorphic vector bundle on X supplied with an involution c_F covering c_X and anti-linear on the fibers, so that it is a real structure on the total space); it is given by $s \mapsto c_F \circ s \circ c_X$. This formula applies as well in a more general situation, when $F \to X$ is a holomorphic fibration with abelian groups as fibers (so that $\mathcal{F}$ is a sheaf of abelian groups) and $c_F\colon F \to F$ is a fiberwise additive real structure covering c_X. Although, in general, $\mathcal{F}$ is not a coherent sheaf, we will still refer to the result as the *canonical real structure* on $\mathcal{F}$.

The canonical real structure on $\mathcal{O}_X$ defines involutions on the other two members of the exponential sequence

$$0 \longrightarrow \mathbb{Z}_X \xrightarrow{2\pi i} \mathcal{O}_X \longrightarrow \mathcal{O}_X^* \longrightarrow 0.$$

Note that the resulting involution on the constant sheaf $\mathbb{Z}_X$ differs from the canonical involution above by (-1). In order to emphasize this nonstandard real structure, we will use the notation $\mathbb{Z}_X^-$ (and, more generally, G_X^-).

2.1.5. Let $\mathcal{A}$ be a sheaf on X with an involution $c\colon \mathcal{A} \to c_X^* \mathcal{A}$. Denote by $\pi\colon X \to X/c_X$ the projection and consider the complex

$$\pi_* \mathcal{A}^*\colon\ 0 \longrightarrow \pi_* \mathcal{A} \xrightarrow{1-c} \pi_* \mathcal{A} \xrightarrow{1+c} \pi_* \mathcal{A} \xrightarrow{1-c} \dots$$

of sheaves on X/c_X (the leftmost copy of $\pi_* \mathcal{A}$ being of degree zero; for simplicity, we use the same notation c for the automorphism $\pi_* c\colon \pi_* \mathcal{A} \to \pi_* \mathcal{A}$). We will refer to the hypercohomology $\mathbf{H}^*(X/c_X; \pi_* \mathcal{A}^*)$ as the hypercohomology of $(\mathcal{A}, c)$ and denote it $\mathbf{H}^*(\mathcal{A}, c)$ (or just $\mathbf{H}^*(\mathcal{A})$, when c is understood). For the constant sheaf G_X with its canonical real structure we will also use the notation $\mathbf{H}^*(X; G)$.

Recall that there are natural spectral sequences

$$H^q(X/c_X; \mathcal{H}^p(\mathbb{Z}_2; \mathcal{A})) \Longrightarrow \mathbf{H}^{p+q}(\mathcal{A}), \tag{2.1.6}$$

where $\mathcal{H}^p(\mathbb{Z}_2;\mathcal{A})$ stand for the cohomology sheaves of $\pi_*\mathcal{A}^*$, and

$$(2.1.7)\qquad H^p(\mathbb{Z}_2;H^q(X/c_X;\pi_*\mathcal{A})) = H^p(\mathbb{Z}_2;H^q(X;\mathcal{A})) \Longrightarrow \mathbf{H}^{p+q}(\mathcal{A}).$$

(Since π is finite-to-one, the higher direct images $R^i\pi_*$, $i > 0$, vanish and one has $H^q(X/c_X;\pi_*\mathcal{A}) = H^q(X;\mathcal{A})$.) Furthermore, since π is finite-to-one, one can calculate $\mathbf{H}^*(\mathcal{A})$ using c_X-invariant Čech resolutions of $\mathcal{A}$. More precisely, given a c_X-invariant open covering $\mathcal{U} = \{U_i\}$ of X, one can consider the bi-complex

$$(\check{C}_{\mathcal{U}}^{p,*}, d_1, d_2) = \bigoplus_{p\geq 0}(\check{C}_{\mathcal{U}}^*(\mathcal{A}), d_2), \qquad d_1 = 1-(-1)^p c\colon\ C_{\mathcal{U}}^{p,*} \to C_{\mathcal{U}}^{p+1,*}$$

(direct sum of copies of the ordinary Čech complex with the first differential given above). Then $\mathbf{H}^n(\mathcal{A})$ is the limit, over all coverings, of the cohomology $H^n(\check{C}_{\mathcal{U}}^{*,*})$.

2.2. Kalinin's spectral sequence. Let X be a a finite dimensional paracompact Hausdorff topological space with an involution $c_X\colon X \to X$.

2.2.1. The *Borel construction* over (X, c_X) is the twisted product

$$X_c = X \times_c S^\infty = (X \times S^\infty)/(x, \mathbf{r}) \sim (c_X(x), -\mathbf{r}).$$

The cohomology groups $H_c^*(X;G) = H^*(X_c;G)$ are called the *equivariant cohomology* of X (with coefficients in an abelian group G). Note that the subscript c stands for the involution $c = c_X$; as we never use cohomology with compact supports, we hope that this notation will not lead to a confusion. The Leray spectral sequence of the fibration $X_c \to \mathbb{R}\mathrm{p}^\infty = S^\infty/\pm\,\mathrm{id}$ with fiber X is called the *Borel-Serre spectral sequence* of (X, c_X):

$${}^2E^{pq}(X;G) = H^p(\mathbb{Z}_2;H^q(X;G)) \Longrightarrow H_c^{p+q}(X;G).$$

Sometimes it is convenient to start the sequence at the term ${}^1E^{pq} = H^q(X;G)$ with the differential ${}^1d^{p*} = 1-(-1)^p c_X^*$.

There is a canonical isomorphism $H_c^p(X;G) = \mathbf{H}^p(X;G)$, and the Borel-Serre spectral sequence is isomorphic to the spectral sequence (2.1.7) for the constant sheaf $\mathcal{A} = G_X$. If G is a commutative ring, then the Borel-Serre spectral sequence is a spectral sequence of $H^*(\mathbb{R}\mathrm{p}^\infty;G)$-algebras.

2.2.2. Let $G = \mathbb{Z}_2$, and let $\hbar \in H^1(\mathbb{R}\mathrm{p}^\infty;\mathbb{Z}_2) = \mathbb{Z}_2$ be the generator. Assume, in addition, that X is a ***CW***-complex of finite dimension. Then the *stabilization homomorphisms*

$$\cup\hbar\colon\ {}^rE^{pq}(X;\mathbb{Z}_2) \to {}^rE^{p+1,q}(X;\mathbb{Z}_2), \quad \cup\hbar\colon\ H_c^n(X;\mathbb{Z}_2) \to H_c^{n+1}(X;\mathbb{Z}_2)$$

are isomorphisms for $p \gg 0$ and one has

$$\lim_{n\to\infty} H_c^n(X;\mathbb{Z}_2) = H^{n\gg 0}(\operatorname{Fix} c_X \times \mathbb{R}\mathrm{p}^\infty;\mathbb{Z}_2) = H^*(\operatorname{Fix} c_X;\mathbb{Z}_2).$$

Thus, one obtains a $\mathbb{Z}$-graded spectral sequence

$${}^rH^q(X;\mathbb{Z}_2) = \lim_{p\to\infty} {}^rE^{pq}(X;\mathbb{Z}_2),$$

called *Kalinin's spectral sequence* of (X, c_X). As above, it is convenient to start the sequence at the term ${}^1H^*(X;\mathbb{Z}_2) = H^*(X;\mathbb{Z}_2)$ with the differential ${}^1d^* = 1 + c^*$.

Kalinin's spectral sequence converges to $H^*(\operatorname{Fix} c_X;\mathbb{Z}_2)$. More precisely, there is an increasing filtration $\{\mathcal{F}_q\} = \{\mathcal{F}_q(X;\mathbb{Z}_2)\}$ on $H^*(\operatorname{Fix} c_X;\mathbb{Z}_2)$, called *Kalinin's filtration*, and homomorphisms

$$\mathrm{b}v^q\colon\ {}^\infty H^q(X;\mathbb{Z}_2) \to H^*(\operatorname{Fix} c_X;\mathbb{Z}_2)/\mathcal{F}_{q-1},$$

called *Viro homomorphisms*, which establish isomorphisms of the graded groups. In general, the convergence does **not** respect the ordinary grading of $H^*(\operatorname{Fix} c_X;\mathbb{Z}_2)$.

The Smith inequality in 2.1.1 can be derived from Kalinin's spectral sequence, and c_X is an M-involution if and only if the sequence degenerates at ${}^1H$. If the sequence degenerates at ${}^2H$, the involution (real variety, etc.) is called $\mathbb{Z}_2$-*Galois maximal*.

A similar construction applies to the homology, producing a $\mathbb{Z}$-graded spectral sequence ${}^rH_q(X;\mathbb{Z}_2)$ starting from $H_q(X;\mathbb{Z}_2)$ and converging to $H_*(\operatorname{Fix} c_X;\mathbb{Z}_2)$. The corresponding decreasing filtration on $H_*(\operatorname{Fix} c_X;\mathbb{Z}_2)$ and Viro homomorphisms are denoted by $\{\mathcal{F}^q(X;\mathbb{Z}_2)\}$ and $\mathrm{b}v_q\colon \mathcal{F}^q \to {}^\infty H_q$, respectively.

2.2.3. The cup-products in $H^*(X;\mathbb{Z}_2)$ descend to a multiplicative structure in ${}^rH^*(X;\mathbb{Z}_2)$, so that ${}^rH^*$ is a $\mathbb{Z}_2$-algebra and the differentials ${}^rd^*$ are differentiations for all $r \geq 2$, i.e., ${}^rd^*(x \cup y) = {}^rd^*x \cup y + x \cup {}^rd^*y$. The filtration $\mathcal{F}_*$ and Viro homomorphisms $\mathrm{b}v^*$ are multiplicative, i.e., $\mathcal{F}_p \cup \mathcal{F}_q \subset \mathcal{F}_{p+q}$ and $\mathrm{b}v^*(x \cup y) = \mathrm{b}v^* x \cup \mathrm{b}v^* y$.

2.2.4. If X is a closed connected n-manifold and $\operatorname{Fix} c_X \neq \varnothing$, Kalinin's spectral sequence inherits Poincaré duality: for each $r \leq \infty$ one has ${}^rH^n(X;\mathbb{Z}_2) = \mathbb{Z}_2$, the cup-product ${}^rH^p(X;\mathbb{Z}_2) \otimes {}^rH^{n-p}(X;\mathbb{Z}_2) \to \mathbb{Z}_2$ is a perfect pairing, and the differentials ${}^rd^p$ and ${}^rd^{n-p-r+1}$, $1 \leq r < \infty$, are dual to each other.

The last member $\mathcal{F}^n$ of the homological filtration is the group $\mathbb{Z}_2$ spanned by the class $w^{-1}(\nu) \cap [X_R]$, where ν is the normal bundle of $X_\mathbb{R}$ in X and $w(\nu)$

is its total Stiefel-Whitney class. (Recall that, if X is a complex manifold and c is a real structure, the normal bundle ν is canonically isomorphic to the tangent bundle τ of $X_{\mathbb{R}}$; the isomorphism is given by the multiplication by i.) Hence, in terms of the cohomology of $\operatorname{Fix} c$ the Poincaré duality above can be stated as follows: the pairing $(x, y) \mapsto \langle x \cup y \cup w^{-1}(\nu), [X_{\mathbb{R}}] \rangle \in \mathbb{Z}_2$ is perfect and, with respect to this pairing, one has $\mathcal{F}_{n-q-1} = \mathcal{F}_q^{\perp}$.

2.2.5. Now, let $G = \mathbb{Z}$, and let $h \in H^2(\mathbb{R}p^\infty; \mathbb{Z}) = \mathbb{Z}_2$ be the generator. Assume, as above, that X is a ***CW***-complex of finite dimension. Then the *stabilization homomorphisms*

$$\cup h\colon\ {}^rE^{pq}(X; \mathbb{Z}) \to {}^rE^{p+2,q}(X; \mathbb{Z}), \quad \cup h\colon\ H_c^n(X; \mathbb{Z}) \to H_c^{n+2}(X; \mathbb{Z})$$

are isomorphisms for $p \gg 0$ and one has

$$\begin{aligned} \lim_{k\to\infty} H_c^{2k}(X; \mathbb{Z}) &= H^{2k\gg 0}(\operatorname{Fix} c_X \times \mathbb{R}p^\infty; \mathbb{Z}) = H^{\mathrm{even}}(\operatorname{Fix} c_X; \mathbb{Z}_2), \\ \lim_{k\to\infty} H_c^{2k+1}(X; \mathbb{Z}) &= H^{2k+1\gg 0}(\operatorname{Fix} c_X \times \mathbb{R}p^\infty; \mathbb{Z}) = H^{\mathrm{odd}}(\operatorname{Fix} c_X; \mathbb{Z}_2). \end{aligned}$$

(We use the notation $H^{p \bmod 2} = \bigoplus_{i=p \bmod 2} H^i$, $H^{\mathrm{even}} = H^{0 \bmod 2}$, $H^{\mathrm{odd}} = H^{1 \bmod 2}$.) Thus, one obtains a $(\mathbb{Z}_2 \times \mathbb{Z})$-graded spectral sequence

$${}^rH^{pq}(X; \mathbb{Z}) = \lim_{k\to\infty} {}^rE^{2k+p,q}(X; \mathbb{Z}), \quad p \in \mathbb{Z}_2,$$

which is also called *Kalinin's spectral sequence* of (X, c_X) (with coefficients in $\mathbb{Z}$). It converges to $H^{\mathrm{even}}(\operatorname{Fix} c_X; \mathbb{Z}_2) \oplus H^{\mathrm{odd}}(\operatorname{Fix} c_X; \mathbb{Z}_2)$, i.e., there are increasing filtrations $\{\mathcal{F}_q^p\} = \{\mathcal{F}_q^p(X; \mathbb{Z})\}$ on $H^{p \bmod 2}(\operatorname{Fix} c_X; \mathbb{Z}_2)$, $p \in \mathbb{Z}_2$, called *Kalinin's filtration*, and homomorphisms $\mathrm{b}v^{pq}\colon\ {}^\infty H^{pq}(X; \mathbb{Z}) \to H^{p \bmod 2}(\operatorname{Fix} c_X; \mathbb{Z}_2)/\mathcal{F}_{q-1}^p$, called *Viro homomorphisms*, which establish isomorphisms of the graded groups.

As in 2.2.2, one can start the sequence at the term ${}^1H^{pq}(X; \mathbb{Z}) = H^q(X; \mathbb{Z})$ with differential ${}^1d^{pq} = 1 - (-1)^{p+q}c^*$. If the sequence ${}^rH^{**}(X; \mathbb{Z})$ degenerates at ${}^2H$, the involution (real variety, etc.) is called *$\mathbb{Z}$-Galois maximal*.

2.2.6. Kalinin's spectral sequence ${}^rH^{**}(X; \mathbb{Z})$ is multiplicative (in the same sense as in 2.2.3), the multiplicative structure inducing the product

$$x \otimes y \mapsto x \cup y + \operatorname{Sq}^1 x \cup \operatorname{Sq}^1 y$$

in the limit term $H^{\mathrm{even}}(\operatorname{Fix} c_X; \mathbb{Z}_2) \oplus H^{\mathrm{odd}}(\operatorname{Fix} c_X; \mathbb{Z}_2)$. (Here $\operatorname{Sq}^1\colon H^p(\,\cdot\,; \mathbb{Z}_2) \to H^{p+1}(\,\cdot\,; \mathbb{Z}_2)$ stands for the Bockstein homomorphism.)

2.2.7. Reduction modulo 2 induces a homomorphism

$$ {}^rH^{0,q}(X;\mathbb{Z}) \oplus {}^rH^{1,q}(X;\mathbb{Z}) \to {}^rH^q(X;\mathbb{Z}_2) $$

of $\mathbb{Z}$-graded spectral sequences, which is compatible with the isomorphism

$$ H^{\mathrm{even}}(\operatorname{Fix} c_X;\mathbb{Z}_2) \oplus H^{\mathrm{odd}}(\operatorname{Fix} c_X;\mathbb{Z}_2) = H^*(\operatorname{Fix} c_X;\mathbb{Z}_2) \xrightarrow{1+\mathrm{Sq}^1} H^*(\operatorname{Fix} c_X;\mathbb{Z}_2) $$

of their limit terms. If $H^*(X;\mathbb{Z})$ is free of 2-torsion, reduction modulo 2 is an isomorphism starting from the term ${}^2H$.

3. Real elliptic surfaces. In what follows, a *surface* is a **nonsingular** complex manifold of complex dimension two. In the few cases when singular surfaces are considered, it is specified explicitly. Proofs of most statements in this section are omitted. We refer the reader to the excellent founding paper by K. Kodaira [Ko], or to the more recent monographs [FM] and [BPV].

3.1. Elliptic surfaces.

3.1.1. An *elliptic surface* is a surface E equipped with an *elliptic fibration*, i.e., a proper holomorphic map $p\colon E \to B$ to a nonsingular curve B (called the *base* of the fibration) such that for all but finitely many points $b \in B$ the fiber $p^{-1}(b)$ is a nonsingular curve of genus 1. We will use the notation $p\colon E|_U \to U$ (or just $E|_U$) for the restriction of the fibration to a subset $U \subset B$. The restriction to the subset $B^\sharp$ of the regular values of p is denoted by $p^\sharp\colon E^\sharp \to B^\sharp$. (In other words, $E^\sharp$ is formed by the nonsingular fibers of p.)

Two fibrations $p\colon E \to B$ and $p'\colon E \to B'$ on the same surface E are considered *identical* if there is an isomorphism $b\colon B \to B'$ such that $p' = p \circ b$. A *morphism* between two elliptic fibrations $p\colon E \to B$ and $p'\colon E' \to B'$ is a pair of proper holomorphic maps $\tilde f\colon E \to E'$ and $f\colon B \to B'$ such that $p' \circ \tilde f = f \circ p$. Two fibrations $p\colon E \to B$ and $p'\colon E' \to B'$ are *isomorphic* if there is a pair of bi-holomorphic maps $\tilde f\colon E \to E'$ and $f\colon B \to B'$ such that $p' \circ \tilde f = f \circ p$.

A compact surface E of positive Kodaira dimension $\kappa(E)$ admits at most one elliptic fibration. All compact surfaces of Kodaira dimension 1 are elliptic; they are called *properly elliptic*.

An *elementary deformation* of elliptic fibrations consists of a nonsingular 3-fold X, a nonsingular surface S, a proper holomorphic map $p\colon X \to S$, and a deformation (in the sense of Kodaira-Spencer) $\pi\colon S \to D = \{z \in \mathbb{C} \mid |z| < 1\}$ such that $p \circ \pi$ is a submersion and each restriction $p_t\colon X_t \to S_t$ of p to the slices $S_t = \pi^{-1}(t)$ and $X_t = p^{-1}S_t$, $t \in D$, is an elliptic fibration. The restrictions p_t are said to be *connected by an elementary deformation*. *Deformation equivalence* of elliptic fibrations is the equivalence relation generated by isomorphisms and elementary deformations. Any deformation of a properly elliptic surface X_0 admits a (unique) structure of deformation of elliptic fibrations.

3.1.2. An elliptic fibration $p\colon E \to B$ is called *real* if both E and B are equipped with real structures $c_E\colon E \to E$ and $c_B\colon B \to B$ so that $c_B \circ p = p \circ c_E$. (When it does not lead to a confusion, we will omit the subscripts in the notation for the real structure.) If E is compact and $\kappa(E) > 0$, then, due to the uniqueness of the elliptic fibration, any real structure $c\colon E \to E$ descends to B.

The notions of morphism, isomorphism, deformation, etc. extend to the real case in the obvious way: one requires that all manifolds involved should be equipped with real structures that are respected by all maps. (For elementary deformations, the real structure on the unit disk $D \subset \mathbb{C}$ is that induced from the complex conjugation.)

3.1.3. In this paper, we only consider *relatively minimal* elliptic fibrations, i.e., those without (-1)-curves in the fibers. For a compact elliptic surface E this is equivalent to the condition $K_E^2 = 0$. As is known, a fiber of an elliptic fibration (as well as any fibration whose generic fiber is a curve of positive genus) can not contain intersecting (-1)-curves. This implies that each elliptic fibration admits a unique relatively minimal model and, in particular, the relatively minimal model is real whenever the original fibration is. Moreover, by Kodaira's results on the stability of exceptional curves, the deformation study of elliptic fibrations (both complex and real) is reduced to the deformation study of their relatively minimal models. In particular, the elliptic fibrations deformation equivalent to a relatively minimal elliptic fibration are relatively minimal.

Any relatively minimal elliptic fibration $p : E \to B$ is *strongly relatively minimal*, i.e., all fiber-to-fiber bi-meromorphic maps $E \to E$ (and, hence, fiber-to-fiber bi-antimeromorphic maps $E \to E$) are regular. In particular, the fibration is uniquely determined by its restriction $p^\sharp\colon E^\sharp \to B^\sharp$ (see 3.1.1), and any real structure on $p^\sharp\colon E^\sharp \to B^\sharp$ extends uniquely to a real structure on $p\colon E \to B$.

3.2. Jacobian surfaces. From now on, we consider only relatively minimal elliptic fibrations **without muptiple fibers**.

3.2.1. To each elliptic fibration $p\colon E \to B$ one can associate its *functional* (or j-) *invariant* $j\colon B \to \mathbb{P}^1$ and its *homological invariant* $R^1p_*\mathbb{Z}_E$.

The functional invariant is the extension to B of the meromorphic function $B^\sharp \to \mathbb{C}$ sending each nonsingular fiber to its j-invariant; following Kodaira, we divide the j-invariant by 12^3, so that its "special" values are $j = 0$ and 1: a nonsingular elliptic curve C with $j(C) = 0$ or 1 has a complex multiplication of order 6 or 4, respectively. Since reversing the complex structure on a nonsingular elliptic curve C transforms $j(C)$ to $\overline{j(C)}$, the functional invariant of a real elliptic fibration is real, $j \circ c_B = \bar{j}$. (When speaking about a real structure on the functional invariant $j\colon B \to \mathbb{P}^1$ we always assume that the real structure on the Riemann sphere $\mathbb{P}^1 = \mathbb{C} \cup \{\infty\}$ is standard, so that the points 0, 1, and ∞ are real.)

An elliptic fibration with $j = \text{const}$ (respectively, $j \neq \text{const}$) is called *isotrivial* (respectively, *nonisotrivial*). In this paper we deal mainly with nonisotrivial fibrations. Note that, unless $j = 0$ or 1, an isotrivial fibration has no singular fibers.

Nonisotrivial fibrations have a *strong extension property*: given a nonsingular curve B, a point $b_0 \in B$, and an elliptic fibration over $B \setminus \{b_0\}$, there is a unique (relatively minimal) elliptic fibration over B whose restriction to $B \setminus \{b_0\}$ is the given one.

3.2.2. The homological invariant (see 3.2.1) is often defined as the monodromy in the 1-cohomology of the nonsingular fibers, i.e., as a local system $\mathcal{M}$ on $B^\sharp$ with fiber $\mathbb{Z} \oplus \mathbb{Z}$, and as such it is just the restriction of $R^1 p_* \mathbb{Z}_E$ to $B^\sharp$. Then $R^1 p_* \mathbb{Z}_E = i_* \mathcal{M}$, i: $B^\sharp \to B$ standing for the inclusion. The homological invariant of a real elliptic fibration inherits a real structure from that on $\mathbb{Z}_E$.

The homological invariant of an elliptic fibration is closely related to its j-invariant. Representing (by means of the elliptic modular function $j(z) = j(\mathbb{C}/L_z)$ where $L_z = \mathbb{Z} + z \cdot \mathbb{Z}$) the complex line $\mathbb{P}^1 \setminus \{\infty\}$ as the quotient of the upper half plane by the modular group $\boldsymbol{PSL}(2, \mathbb{Z}) = \boldsymbol{SL}(2, \mathbb{Z})/\{\pm 1\}$, one equips the space $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ with a $\boldsymbol{PSL}(2, \mathbb{Z})$-principal bundle $\mathcal{P}$. (Here 0 and 1 are the images of the two unstable points of the action.) A local system $\mathcal{M}$ as above is said to *belong to* a holomorphic map j: $B \to \mathbb{P}^1$ if the principal $\boldsymbol{PSL}(2, \mathbb{Z})$-bundle associated with $\mathcal{M}$ is $j^*\mathcal{P}$. The homological invariant of an elliptic fibration belongs to its functional invariant.

Two real structures on a holomorphic map j: $B \to \mathbb{P}^1$ and a $(\mathbb{Z} \times \mathbb{Z})$-local system $\mathcal{M}$ belonging to j are called *concordant* if they are both lifts of the same real structure c_B on B. This is the case if j and $\mathcal{M}$ are, respectively, the functional and homological invariant of a real elliptic fibration.

In both complex and real cases, the passage from j to $\mathcal{M}$ involves a choice of one of the two lifts over each loop $\gamma \subset B^\sharp$; the lifts differ by the multiplication by $-\text{id} \in \boldsymbol{SL}(2, \mathbb{Z})$. Next statement asserts that the elliptic fibrations obtained from distinct lifts differ topologically.

3.2.3 LEMMA. *A matrix $A \in \boldsymbol{SL}(2, \mathbb{Z})$ is never conjugate to $-A$.*

Proof. In fact, the statement holds for the bigger group $\boldsymbol{SL}(2, \mathbb{R})$. If a (2×2)-matrix A is similar to $-A$ and $\det A = 1$, one can easily see that the eigenvalues of A must be $\pm i$, i.e., A is a complex structure (or the rotation through $\pm\pi/2$). Then $-A$ is the rotation in the opposite direction; it is not conjugate to A by an orientation preserving transformation. □

3.2.4. Among all elliptic fibrations with given functional and homological invariants there is a unique, up to isomorphism, elliptic fibration p: $J \to B$ with a section. We equip J with a distinguished section s: $B \to J$ and call the pair (J, s) the *Jacobian elliptic fibration* associated to p: $E \to B$ (or to the given pair of functional and homological invariants). The group of automorphisms of J preserving s is a cyclic group of order 2, 4, or 6, the last two cases occurring only if $j = \text{const}$. In all cases the element of order 2 acts in each nonsingular fiber as the multiplication by (-1).

An elementary deformation

$$p\colon X \to S, \quad \pi\colon S \to D$$

of elliptic fibrations is called *Jacobian*, or a *deformation through Jacobian fibrations*, if it is equipped with a section $s\colon D \to S$ of $p \circ \pi$. This notion extends to the real case in the usual way: one requires that the deformation and the section should be real.

In order to construct the Jacobian elliptic fibration $J = J(E)$ associated to $p\colon E \to B$, one can start from $p^\sharp\colon E^\sharp \to B^\sharp$ and replace each fiber $F_b = p^{-1}(b)$ by its Jacobian $\operatorname{Pic}^0(F_b)$. Then it remains to complete the elliptic fibration $\operatorname{Pic}^0_{B^\sharp} = \bigcup_{b\in B^\sharp} \operatorname{Pic}^0(F_b) \to B^\sharp$ and to take its relatively minimal model. (Note that the completion step requires, in fact, a thorough understanding of singular fibers and their local models. It turns out that the Jacobian fibration has singular fibers of the same types as the original one.) The strong relative minimality implies uniqueness. Moreover, it shows that the construction is functorial, i.e., any (anti-)isomorphism $E \to E'$ of elliptic fibrations induces an (anti-)isomorphism $J(E) \to J(E')$ of their Jacobians respecting the distinguished sections. In particular, the Jacobian elliptic fibration associated to a real elliptic fibration inherits an *associated Jacobian real structure*, namely, the structure $c_J\colon J(E) \to J(E)$ induced by the action of c_E on the Jacobians of the nonsingular fibers. Unless $j = \text{const}$, the only other real structure preserving the section is $-c_J$, i.e., the composition of c_J and the fiberwise multiplication by (-1). The real structures c_J and $-c_J$ are called *opposite* to each other.

A deformation of elliptic surfaces gives rise to a natural deformation of the associated Jacobian surfaces. In particular, if the original deformation is real, so is the resulting Jacobian deformation.

3.2.5. As it follows from the strong extension property (see 3.2.1), for each pair $(j, \mathcal{M})$ consisting of a nonconstant holomorphic map $j\colon B \to \mathbb{P}^1$ and belonging to it $(\mathbb{Z} \times \mathbb{Z})$-local system $\mathcal{M}$ on $B^\sharp = j^{-1}(\{0, 1, \infty\})$ there exist a Jacobian fibration $J(j, \mathcal{M})$ whose functional and homological invariants are j and $\mathcal{M}$. The set of local systems belonging to a given map j is an affine space over $H^1(B^\sharp; \mathbb{Z}_2)$. In particular, their number and, thus, the number of Jacobian fibrations with given functional invariant is 2^r, where $r = b_1(B^\sharp)$. We recall that the monodromy of $\mathcal{M}$ along a small loop about a point $b_0 \in B \smallsetminus B^\sharp$ determines and is determined by the topology of the singular fiber at b_0. Thus, if the types of the singular fibers are fixed (e.g., under the assumption that the fibration is *almost generic*, i.e., does not have singular fibers other than those of type I_1), the admissible homological invariants form an affine space over $H^1(B;\mathbb{Z}_2)$. (Note, though, that the latter assumption imposes certain restrictions to the functional invariant, see 3.3.11 below.)

3.2.6. Let j and $\mathcal{M}$ be as in 3.2.5. Assume that $B_\mathbb{R} \neq \varnothing$. If j and $\mathcal{M}$ are equipped with concordant real structures, then, as it follows from the strong relative minimality and the local Torelli theorem for elliptic curves with a marked point, the real structures lift to a real structure on $J(j,\mathcal{M})$ that makes it a real Jacobian fibration with given real invariants j and $\mathcal{M}$. When nonempty, the set of real local systems belonging to and concordant with a given real map $j\colon B \to \mathbb{P}^1$ is an affine space over $\mathbf{H}^1(B^\sharp;\mathbb{Z}_2)$ (see 4.1.7; the existence question is discussed in 3.3.9). Therefore, their number and, thus, the number of real Jacobian fibrations with a given real functional invariant j is equal to 2^r, where $r = \dim\,(H^1(B^\sharp;\mathbb{Z}_2))^c + 1$, see 4.1.8. (Note that the term $1 = \dim H^1(\mathbb{Z}_2;H^0(B^\sharp;\mathbb{Z}_2))$ accounts for the two distinct real structures on a given complex Jacobian fibration, see 3.2.4, and 2^{r-1} is the number of local systems admitting a concordant real structure.) As in 3.2.5, if the fibration is assumed almost generic, this number reduces to 2^s, where $s = \dim\,(H^1(B;\mathbb{Z}_2))^c + 1$.

3.2.7. Removing from a Jacobian fibration J the singular points (including multiple components) of its singular fibers, one obtains an analytic family $J^{\mathrm{ab}} \to B$ of (not necessarily connected) abelian Lie groups with s as the zero section. Any section of J is contained in J^{ab}, and any two sections differ by a *translation*, i.e., an automorphism of $p\colon J \to B$ which is a translation in each nonsingular fiber. Furthermore, one can introduce two sheafs of abelian groups: the sheaf $\tilde{\mathcal{J}}$ of germs of holomorphic sections of J^{ab} and the sheaf $\mathcal{J}$ of germs of sections of J^{ab} intersecting each fiber at the same component as s. The two sheafs differ by a skyscraper $\mathcal{S}$ having finite fibers and concentrated at the points of B corresponding to reducible singular fibers with at least two simple components,

$$0 \longrightarrow \mathcal{J} \longrightarrow \tilde{\mathcal{J}} \longrightarrow \mathcal{S} \longrightarrow 0. \tag{3.2.8}$$

(In fact, the order of the stalk $\mathcal{S}_b$ at a point $b \in B$ is exactly the number of simple components in the fiber $p^{-1}(b)$.) From the exponential sequence over J one can also obtain the following short exact sequence for $\mathcal{J}$:

$$0 \longrightarrow R^1p_*\mathbb{Z}_J \xrightarrow{2\pi i} R^1p_*\mathcal{O}_J \longrightarrow \mathcal{J} \longrightarrow 0. \tag{3.2.9}$$

Recall that, for any elliptic fibration $p\colon E \to B$, the sheaf $R^1p_*\mathcal{O}_E$ is of the form $\mathcal{O}_B(L^{-1})$, where L is a certain line bundle on B determined solely by the functional and homological invariants of the fibration. If the fibration has a section $s\colon B \to E$, then L^{-1} is the normal bundle of s; its degree (i.e., the self-intersection of s) is negative unless $j = \mathrm{const}$.

3.2.10. Let $p\colon E \to B$ be an elliptic fibration and $J \to B$ its Jacobian. Any (anti-)automorphism $g\colon E \to E$ induces an (anti-)automorphism $J(g)\colon J \to J$ preserving the distinguished section (see 3.2.4). The kernel $\operatorname{Aut}_0 E$ of the map $g \mapsto J(g)$ is formed by the translations of E. Obviously, there is a canonical isomorphism $\operatorname{Aut}_0 E = \operatorname{Aut}_0 J$ and both of the groups are isomorphic to the *Mordel-Weil group* $\Gamma(B; \tilde{\mathcal{J}})$ (the group of sections of $J \to B$). Furthermore, if E has a section itself, it is isomorphic to J, and the set of isomorphisms $\varphi\colon E \to J$ identical on the Jacobian is an affine space over $\Gamma(B; \tilde{\mathcal{J}})$: one has $\varphi + t = \varphi \circ t_E = t_J \circ \varphi$, where $t \in \Gamma(B; \tilde{\mathcal{J}})$ is a section and t_E, t_J are the corresponding translations of E and J, respectively.

Recall that the Mordel-Weil group of any compact nonisotrivial elliptic fibration is discrete. This follows, e.g., from the fact that the line bundle L has no sections (as it has negative degree).

Now, let c_E be a real structure on E and c_J the Jacobian real structure on J. The latter induces a real structure c on $\tilde{\mathcal{J}}$, $s \mapsto c_J \circ s \circ c_B$, cf. 2.1.3, which is compatible with (3.2.8) and (3.2.9). The set of all real structures on E whose Jacobian is c_J is an affine space over the subgroup $\operatorname{Ker}(1 + c) \subset \Gamma(B; \tilde{\mathcal{J}})$, the affine action being $c_E + t = t_E \circ c_E$. (Here, the condition $(1 + c)t = 0$ is necessary and sufficient for the composition $t_E \circ c_E$ to be an involution.) The shift t_E by a section $t \in \Gamma(B; \tilde{\mathcal{J}})$ is real (i.e., commutes with c_E) if and only if $(1 - c)t = 0$. Note that the condition does not depend on c_E; thus, t_E commutes with any real structure whose Jacobian is c_J. More generally, for any such real structure c_E on E one has $t_E^{-1} \circ c_E \circ t_E = c_E - (1 - c)t$. In particular, the set of all real structures on E with the given Jacobian c_J is an affine space over $H^1(\mathbb{Z}_2; \Gamma(B; \tilde{\mathcal{J}}))$.

3.3. Trigonal curves and Weierstraß models. Traditionally, elliptic curves with a rational point are described via the so called Weierstraß equation. In the case of elliptic surfaces, this approach leads to trigonal curves on ruled surfaces.

3.3.1. Trigonal curves. Let $q\colon \Sigma \to B$ be a geometrically ruled surface with a distinguished section s. A *trigonal curve* on Σ is a reduced curve $C \subset \Sigma$ disjoint from s and such that the restriction $q\colon C \to B$ is of degree three. Given a trigonal curve $C \subset \Sigma$, the fiberwise center of gravity of the three points of C (regarded as points in the affine fiber of $\Sigma \setminus s$) defines an additional section 0 of Σ; thus, the 2-bundle whose projectivization is Σ splits and, after a renormalization, can be chosen in the form $1 \oplus Y$. We choose the normalization so that the projectivization of the Y summand is the zero section.

Any trigonal curve can be given by a *Weierstraß equation*; in appropriate affine charts it has the form

$$x^3 + g_2 x + g_3 = 0, \tag{3.3.2}$$

where g_2 and g_3 are certain sections of Y^2 and Y^3, respectively, and x is a coordinate such that $x = 0$ is the zero section and $x = \infty$ is the distinguished section s. The sections g_2, g_3 are determined by the curve uniquely up to the transformation

$$(g_2, g_3) \mapsto (t^2 g_2, t^3 g_3), \quad t \in \Gamma(B, \mathcal{O}_B^*).$$

If both (Σ, s) and C are real, then Y is a real line bundle and the sections g_2, g_3 can also be chosen real; they are defined uniquely up to the above transformation with a real section t.

The *j-invariant* of a trigonal curve $C \subset \Sigma$ is the function $j\colon B \to \mathbb{P}^1$ given by

$$j = \frac{4g_2^3}{\Delta}, \quad \Delta = 4g_2^3 + 27g_3^2. \tag{3.3.3}$$

Geometrically, the value of j at a generic point $b \in B$ is the usual j-invariant of the quadruple of points cut by the union $C \cup s$ in the projective line $q^{-1}(b)$.

As the equation suggests, in general the j-invariant does not change continuously when the curve is deformed; even the degree of j can change.

From now on, by a *deformation* of a trigonal curve $C \subset \Sigma$ we mean a deformation of the quadruple (B, q, s, C) (i.e., neither Σ nor B are assumed fixed). As usual, the *deformation equivalence* of trigonal curves is the equivalence relation generated by deformations and isomorphisms (of the quadruples as above).

A trigonal curve $C \subset \Sigma$ is called *almost generic* if it is nonsingular and has no vertical flexes. If this is the case, the j-invariant $j\colon B \to \mathbb{P}^1$ has degree $\deg j = 6d$, where $d = \deg Y$, the point $\infty \in \mathbb{P}^1$ is a regular value of j, its $6d$ pull-backs corresponding to the vertical tangents of the curve, and all pull-backs of 0 and 1 have ramification index 0 mod 3 (respectively, 0 mod 2). By an arbitrary small deformation (including a change of the complex structure of the base) one can achieve that the j-invariant have so called *generic branching behavior* (which, in fact, is highly non-generic for a function $B \to \mathbb{P}^1$): in addition to the above conditions one requires that the ramification index of each pull-back of 0 (respectively, 1) should be exactly 3 (respectively, 2). (Note that other critical values, which j unavoidably has, are irrelevant.) A trigonal curve whose j-invariant has generic branching behavior is called *generic*.

3.3.4. Topology. The real part $\Sigma_{\mathbb{R}}$ of a real geometrically ruled surface $q\colon \Sigma \to B$ consists of several connected components Σ_i, one over each component B_i of $B_{\mathbb{R}}$. Each Σ_i is either a torus or a Klein bottle. If Σ has the form $\mathbb{P}(1 \oplus Y)$ as above, then a component Σ_i is orientable if and only if the restriction Y_i of the real part $Y_{\mathbb{R}}$ of Y to the corresponding component B_i is topologically trivial. In other words, if Y is given by a real divisor D, then a component Σ_i is orientable if and only if the degree $\deg(D \cap B_i)$ is even. The latter remark shows also that $\sum_i \deg Y_i = \deg Y \bmod 2$. Hence, $\Sigma_{\mathbb{R}}$ is necessarily nonorientable whenever $\deg Y$ is odd.

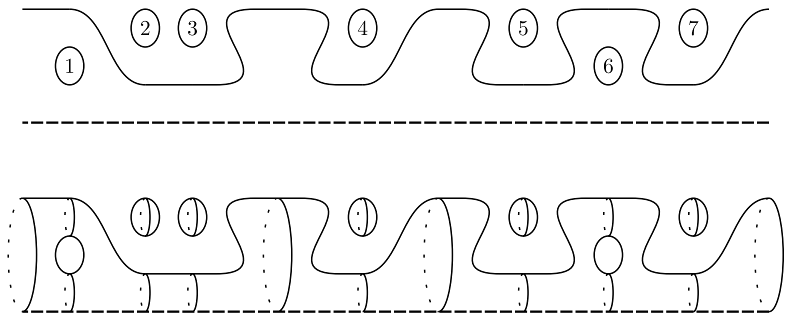


Figure 1. A typical nonhyperbolic trigonal curve (top) and a corresponding Jacobian surface (bottom); the horizontal dotted lines represent the distinguished sections s of the surfaces.

Let $C \subset \Sigma$ be a real trigonal curve, and let $q_{\mathbb{R}}\colon C_{\mathbb{R}} \to B_{\mathbb{R}}$ be the projection. The real part $C_{\mathbb{R}}$ splits into groups of components $C_i = q_{\mathbb{R}}^{-1}(B_i)$. Each restriction $q_{\mathbb{R}}\colon C_i \to B_i$ is onto. A component B_i of $B_{\mathbb{R}}$ (and the corresponding group C_i) is called *hyperbolic* if the restriction $q_{\mathbb{R}}\colon C_i \to B_i$ is generically three-to-one; otherwise, it is called *nonhyperbolic*. A curve C with non-empty real part is called *hyperbolic* if all its groups C_i are hyperbolic; otherwise, it is called *nonhyperbolic*.

Over a hyperbolic component B_i, the group C_i of an almost generic curve consists of a 'central' component which projects to B_i homeomorphically and two (if Σ_i is orientable) or one (if Σ_i is nonorientable) additional components; the restriction of $q_{\mathbb{R}}$ to the union of the additional components is a double covering, trivial in the former case and nontrivial in the latter case.

The group C_i of an almost generic curve over a nonhyperbolic component B_i looks as shown in Figure 1. More precisely, C_i has a "long" component mapped onto B_i and several contractible components, commonly called ovals; the long component may contain a few "zigzags", which are also preserved by fiberwise isotopies. For the purpose of this paper, we define *ovals* and *zigzags* as the connected components of the set $\{b \in B_i | \# q_{\mathbb{R}}^{-1} \geq 2\}$; ovals are those whose pull-back is disconnected. The set of all ovals within a nonhyperbolic component B_i inherits from B_i a pair of opposite cyclic orders.

Pick a nonhyperbolic component B_i. Fix a section $\sigma\colon B_i \to \Sigma_i$ disjoint from s and taking each oval inside the corresponding contractible component; such a section is unique up to homotopy. A set of consecutive ovals in B_i is called a *chain* if between any two neighboring ovals of the set the section σ intersects the long component an even number of times. For example, in Figure 1 the maximal chains are [1], [2, 3, 4, 5], [6], and [7] (assuming that Σ_i is orientable; otherwise, ovals 7 and 1 form a single chain). A chain of ovals is called *complete* if it contains all ovals in a single component B_i.

The notions of oval and chain extend to all nonsingular trigonal curves. Note that a nonhyperbolic nonsingular curve cannot be isotrivial; hence, it can be perturbed to an almost generic one.

3.3.5. Weierstraß models. The *Weierstraß model* of a Jacobian elliptic surface $p\colon J \to B$ is obtained from J by contracting all components of the fibers of p that do not intersect the distinguished section s. The result is a proper map $p\colon J^{\mathrm{w}} \to B$, where J^{w} has at worst simple singular points, and a section $s\colon B \to J^{\mathrm{w}}$ not passing through the singular points of J^{w}. The original Jacobian surface J is recovered from J^{w} by resolving its singularities.

The quotient of J^{w} by the fiberwise multiplication by (-1) is a geometrically ruled surface Σ over B, the section s mapping to a section of Σ. The projection $J^{\mathrm{w}} \to \Sigma$ is the double covering defined by the (fiberwise) linear system $|2s|$ on J^{w}; its branch curve is the disjoint union of the exceptional section s and a certain trigonal curve C on Σ. In particular, Σ has the form $\mathbb{P}(1 \oplus Y)$ and $Y = L^2$, where L is the conormal bundle of s in J (cf. 3.2.7).

The sections g_2, g_3 defining C, see (3.3.2), must satisfy the following conditions:

(1) the discriminant $\Delta = 4g_2^3 + 27g_3^2$ is not identically zero, and

(2) at each point $b \in B$ one has $\min(3\operatorname{ord}_b(g_2), 2\operatorname{ord}_b(g_3)) < 12$.

(The former condition ensures that generic fibers are nonsingular elliptic curves, and the latter implies that all singular points of J^{w} are simple.) Conversely, given a ruled surface $\Sigma = \mathbb{P}(1 \oplus Y)$ with a section s and a trigonal curve (3.3.2), a choice of a square root L of Y defines a unique double covering of Σ ramified at s and the curve; if the pair (g_2, g_3) in (3.3.2) satisfies (1) and (2) above, the double covering is the Weierstraß model of a Jacobian elliptic surface. The j-invariant of the resulting surface is given by (3.3.3).

3.3.6. Real structures. The construction above is natural. Hence, if the Jacobian surface (J, s) is real, so are the ruled surface Σ, its section s, and the branch curve C. Conversely, if the surface (Σ, s), curve C, and square root L are real, then the Jacobian surface J resulting from the construction above inherits two opposite real structures. Under the assumptions, $\Sigma_{\mathbb{R}}$ splits into two *halves* with common boundary $C_{\mathbb{R}} \cup s_{\mathbb{R}}$, the halves being the projections of the real parts of the two real structures on J. A choice of one of the two real structures is equivalent to a choice of one of the two halves.

The real part $J_{\mathbb{R}}$ is a double of the corresponding half; it looks as shown in Figure 1. It splits into groups of components $J_i = p_{\mathbb{R}}^{-1}(B_i)$. Each group J_i has a distinguished component that contains the section s over B_i; we call this component *principal*. Its orientability is governed by the restriction L_i of the real part $L_{\mathbb{R}}$ of L to B_i: the principal component is orientable if and only if L_i is topologically trivial. Besides, there are a few *extra* components disjoint from the section $s_{\mathbb{R}}$. In the nonhyperbolic case all extra components are spheres. In the hyperbolic case, there is exactly one extra component, which is either a torus or a Klein bottle, depending on whether L_i is trivial or not. Thus, in the hyperbolic case the two components are either both orientable or both not.

The following lemma states that a real Jacobian surface and its branch curve have the same discrepancy.

3.3.7 LEMMA. *An almost generic Jacobian surface J is an $(M-d)$-variety if and only if so is the trigonal part of the branch curve of the Weierstraß model of J.*

Proof. Indeed, the isomorphism $\pi_1(J) \to \pi_1(B)$, see [FM, Proposition 2.2.1], gives $\dim H_1(J;\mathbb{Z}_2) = \dim H_1(B;\mathbb{Z}_2)$; the other Betti numbers are controlled using the Riemann-Hurwitz formula $\chi(J) = 2\chi(\Sigma) - \chi(C) = 4\chi(B) - \chi(C)$ and Poincaré duality. The Betti numbers of the real part $J_{\mathbb{R}}$ are found using the description of its topology given in 3.3.6, and the statement follows from a simple comparison. □

3.3.8. The homological invariant. The Weierstraß model gives a clear geometric interpretation of the $\boldsymbol{PSL}(2,\mathbb{Z})$-bundle $j^*\mathcal{P}$ defined by j, see 3.2.2. Indeed, the modular group $\boldsymbol{PSL}(2,\mathbb{Z})$ is naturally identified with the factorized braid group B_3/Δ^2, which, in turn, can be regarded as the mapping class group of the triad $(F; s\cap F, C\cap F)$, where F is a generic fiber of the ruling. Thus, $j^*\mathcal{P}$ is merely the monodromy $\pi_1(B^\sharp) \to B_3/\Delta^2$ of the trigonal curve.

As explained in 3.2.5, the homological invariants belonging to a given functional invariant j form an affine space over $H^1(B^\sharp, \mathbb{Z}_2)$. The branch curve C narrows this choice down to $H^1(B, \mathbb{Z}_2)$, as its singularities and vertical tangents determine the singular fibers of the covering elliptic surfaces. (Roughly speaking, the singularities of C are encoded, in addition to j, in the presence and multiplicities of the common roots of the sections g_2, g_3). The rest of the homological invariant is recovered via the choice of the square root L of Y.

In the real case, the choice is narrowed down to $\mathbf{H}^1(B;\mathbb{Z}_2)$, see 3.2.6, and partially it can be made canonical using the correspondence $L \mapsto \bigoplus_i w_1(L_i)$, which is an affine map from the set of homological invariants onto the subset

$$\{\alpha \in H^1(B_{\mathbb{R}};\mathbb{Z}_2) \mid \alpha[B_{\mathbb{R}}] = \deg L \bmod 2\}.$$

In other words, the real Jacobian elliptic surfaces with a given branch curve are partially distinguished by the orientability of their principal components.

3.3.9. Existence of the roots. Obviously, the necessary and sufficient condition for the existence of a square root L of a line bundle Y on B (and, hence, the existence of an elliptic surface over a given ruled surface) is that $\deg Y$ should be even.

In the real case, if $B_{\mathbb{R}} \neq \varnothing$, for the existence of a real square root L of a real line bundle Y one should require, in addition, that the real part $Y_{\mathbb{R}}$ is topologically trivial. Indeed, the condition is obviously necessary. For the sufficiency notice that, if $B_{\mathbb{R}} \neq \varnothing$, a line bundle is real if and only if its class in $\operatorname{Pic} B$ is fixed by the induced real structure. In particular, the property to have real roots is invariant under equivariant deformations of Y. On the other hand, the real part $\operatorname{Pic}^0_{\mathbb{R}} B$ has $2^{b_0(B_{\mathbb{R}})-1}$ connected components which are distinguished by the restrictions to the components B_i (as the real structure on $\operatorname{Pic}^0 B$ is essentially the same as the induced involution in $H^1(B;\mathbb{Z})$). Hence, all bundles of a given degree whose real part is trivial are deformation equivalent.

Combining the above statement with the beginning of 3.3.4, one arrives at the following criteria.

3.3.10 Corollary. *Assume that a real ruled surface* (Σ, s), $\Sigma_{\mathbb{R}} \neq \varnothing$, *and a real trigonal curve* $C \subset \Sigma$ *do define a Jacobian surface. The latter can be chosen real if and only if* $\Sigma_{\mathbb{R}}$ *is orientable.* □

3.3.11. An application: generic surfaces. Any nonisotrivial Jacobian surface can be deformed through Jacobian surfaces to an almost generic one. (In general, that would change the base of the fibration. The deformation can be chosen elementary and arbitrary small.) The j-invariant of an almost generic surface is similar to that of an almost generic curve, the pull-backs of ∞, 0, and 1 corresponding to the singular fibers (of type I_1), fibers with complex multiplication of order 6, and those with complex multiplication of order 4, respectively. By another arbitrary small deformation one can make the surface *generic*, i.e., achieve that the j-invariant have generic branching behavior.

If (J, s) is real, the above deformations can also be chosen real. For proof one can use the same "cut-and-paste" arguments as in the complex case, constructing local deformations and patching them together. To construct a local perturbation, one can use the Weierstraß equation (3.3.2), which contains a versal deformation of the special point (i.e., singular point, vertical flex, or multiple root of g_2 or g_3) of the Weierstraß model of J. After a covering parameter change, any perturbation of J admits a simultaneous resolution of singularities to which one can extend any real structure and any automorphism of the original perturbation. (This follows, e.g., from the Grothendieck-Brieskorn model.) Due to the versality, the covering deformation can be realized as a deformation of Weierstraß models.

If an elliptic fibration is deformed through almost generic ones, its functional invariant does change continuously. Conversely, if an analytic family of degree $12d$ functions having generic branching behavior includes the j-invariant of a Jacobian elliptic fibration J, it results in a unique deformation of J through generic Jacobian fibrations. Observing that Kodaira's proof respects real structures (or using the uniqueness of the deformation), one obtains a real version of the statement: if the fibration J and the family of functions are real, the resulting deformation is real.

4. Real Tate-Shafarevich group. Fix a real Jacobian fibration $p\colon J \to B$, not necessarily compact, with a real section $s\colon B \to J$. The *real* (analytic) *Tate-Shafarevich group* of J is the set $\mathbb{R}Ш = \mathbb{R}Ш(J)$ of isomorphism classes of triples (E, c, φ), where (E, c) is a real elliptic surface (over B) without multiple fibers and φ: $J(E) \to J$ is a real isomorphism. (The group structure on $\mathbb{R}Ш(J)$ is given by Theorem 4.1.1 below.) Our principal result in this section is the fact that the 'discrete part' $\mathbb{R}Ш^{\mathrm{top}} = \mathbb{R}Ш/\mathbb{R}Ш^0$ (where $\mathbb{R}Ш^0$ is the component of unity) is a topological invariant of the pair (p, c_J).

4.1. Topological invariance. Let p: $J \to B$ be as above and let c be the canonical real structure on the sheaf $\tilde{\mathcal{J}}$ of germs of holomorphic sections of p, see 3.2.10.

4.1.1 THEOREM. *There is a natural isomorphism* $\mathbb{R}Ш(J) = \mathbf{H}^1(\tilde{\mathcal{J}}, c)$.

Proof. We mimic the standard proof of the similar result for the complex (analytic) Tate-Shafarevich group $Ш(J)$. Pick a triple $(E, c, \varphi) \in \mathbb{R}Ш(J)$ and use φ to identify the Jacobian of E and J. Since E has no multiple fibers, one can cover B by c_B-invariant open sets U_i so that each restriction $E_i = E|_{U_i}$ has a section (not necessarily real), or, equivalently, there is an isomorphism φ_i: $E_i \to J_i = J|_{U_i}$. Let $c_i = \varphi_i^{-1} \circ c_J \circ \varphi_i$ be the real structure on E_i induced by φ_i. Then the restriction of c_E to E_i has the form $c_i + s_i$ for some section $s_i \in \Gamma(U_i; \tilde{\mathcal{J}})$ satisfying $(1 + c)s_i = 0$, see 3.2.10. The restrictions of φ_i and φ^{-1} to the intersection $U_{ij} = U_i \cap U_j$ have the same Jacobian and, hence, differ by a section $t_{ij} \in \Gamma(U_{ij}; \tilde{\mathcal{J}})$: one has $\varphi_j = \varphi_i + t_{ij}$ (see 3.2.10 again) and, as usual, the 1-cochain (t_{ij}) must be a cocycle in the Čech complex $(\check{C}^*_{\mathcal{U}}(\tilde{\mathcal{J}}), d_2)$. Finally, the real structures on $c_i + s_i$ and $c_j + s_j$ must coincide on $E|_{U_{ij}}$ and, since $c_j = c_i - (1 - c)t_{ij}$ over U_{ij}, one has $s_j - s_i = (1 - c)t_{ij}$. Thus, the sections (s_i, t_{ij}) form a 1-cocycle in the Čech bi-complex $\check{C}^{**}_{\mathcal{U}}(\tilde{\mathcal{J}})$ corresponding to the covering $\mathcal{U} = \{U_i\}$. Any other set of isomorphisms φ'_i: $E_i \to J_i$ differs from φ_i by sections $r_i \in \Gamma(U_i; \tilde{\mathcal{J}})$, $\varphi'_i = \varphi_i + r_i$, and for the new sections s'_i, t'_{ij} one has $t'_{ij} = t_{ij} + r_j - r_i$ and $s'_i = s_i - (1 - c)r_i$. Thus, the new cocycle (s'_i, t'_{ij}) differs from (s_i, t_{ij}) by the coboundary of the 0-cochain (r_i) (and, *vice versa*, changing the cocycle (s_i, t_{ij}) by the coboundary of (r_i) can be realized by replacing the isomorphisms φ_i with $\varphi_i + r_i$).

Conversely, let (s_i, t_{ij}) be a 1-cocycle. Since (t_{ij}) is a 1-cocycle in the ordinary Čech complex, it defines a complex elliptic surface E (by gluing the pieces J_i along their intersections *via* the translations t_{ij}). Since $(1 + c)s_i = 0$, the anti-automorphisms $c_J + s_i$ are real structures on J_i, and the cocycle condition guarantees that these real structures agree on the intersections, thus blending into a real structure on E. □

In addition to the sheaves $\tilde{\mathcal{J}}$, $\mathcal{J}$, $R^1 p_* \mathcal{O}_J$ and exact sequences (3.2.8), (3.2.9) consider the sheaves $\tilde{\mathcal{J}}^{\mathrm{top}}$, $\mathcal{J}^{\mathrm{top}}$, $(R^1 p_* \mathcal{O}_J)^{\mathrm{top}}$ of continuous sections of the corresponding bundles/fibrations and exact sequences

$$0 \longrightarrow \mathcal{J}^{\mathrm{top}} \longrightarrow \tilde{\mathcal{J}}^{\mathrm{top}} \longrightarrow \mathcal{S} \longrightarrow 0, \tag{4.1.2}$$

$$0 \longrightarrow R^1 p_* \mathbb{Z}_J^- \longrightarrow (R^1 p_* \mathcal{O}_J)^{\mathrm{top}} \longrightarrow \mathcal{J}^{\mathrm{top}} \longrightarrow 0. \tag{4.1.3}$$

(For the discrete sheaves $R^1p_*\mathbb{Z}_J^-$ and $\mathcal{S}$ one would have $(R^1p_*\mathbb{Z}_J^-)^{\mathrm{top}} = R^1p_*\mathbb{Z}_J^-$ and $\mathcal{S}^{\mathrm{top}} = \mathcal{S}$.)

4.1.4 LEMMA. *If $\mathcal{A}$ is a sheaf of $\mathbb{R}$-vector spaces and the involution $c\colon \mathcal{A} \to c_B^*\mathcal{A}$ is $\mathbb{R}$-linear, then* $\mathbf{H}^*(\mathcal{A}) = (H^*(B;\mathcal{A}))^c$ (the subspace of c-invariant classes). *In particular, $\mathbf{H}^i(R^1p_*\mathcal{O}_J) = 0$ for $i > 1$, and $\mathbf{H}^i((R^1p_*\mathcal{O}_J)^{\mathrm{top}}) = 0$ for $i > 0$.*

Proof. For any involution c on a vector space U over a field of characteristic 0 the sequence $U \xrightarrow{1-c} U \xrightarrow{1+c} U$ is exact, and the first statement follows from (2.1.7). The rest is immediate. □

4.1.5 COROLLARY. *There is a natural exact sequence* (induced from (3.2.9))

$$\mathbf{H}^1(R^1p_*\mathbb{Z}_J^-) \longrightarrow (H^1(B;R^1p_*\mathcal{O}_J))^c \longrightarrow \mathbf{H}^1(\mathcal{J}) \longrightarrow \mathbf{H}^2(R^1p_*\mathbb{Z}_J^-) \longrightarrow 0.$$

In particular, the discrete part of $\mathbf{H}^1(\mathcal{J})$ is canonically isomorphic to $\mathbf{H}^2(R^1p_*\mathbb{Z}_J^-)$. Furthermore, there are natural isomorphisms $\mathbf{H}^i(\mathcal{J}) = \mathbf{H}^{i+1}(R^1p_*\mathbb{Z}_J^-)$, $i > 1$, and $\mathbf{H}^i(\mathcal{J}^{\mathrm{top}}) = \mathbf{H}^{i+1}(R^1p_*\mathbb{Z}_J^-)$, $i > 0$. □

4.1.6 THEOREM. *There is a canonical isomorphism $\mathbb{R}Ш^{\mathrm{top}}(J) = \mathbf{H}^1(\tilde{\mathcal{J}}^{\mathrm{top}})$ and a natural* (with respect to real fiberwise homeomorphisms) *exact sequence*

$$\mathbf{H}^0(\mathcal{S}) \longrightarrow \mathbf{H}^2(R^1p_*\mathbb{Z}_J^-) \longrightarrow \mathbb{R}Ш^{\mathrm{top}}(J) \longrightarrow \mathbf{H}^1(\mathcal{S}) \longrightarrow \mathbf{H}^3(R^1p_*\mathbb{Z}_J^-).$$

In particular, $\mathbb{R}Ш^{\mathrm{top}}(J)$ is a topological invariant of the pair (p, c_J).

Proof. In view of Theorem 4.1.1, it suffices to show that the group $\mathbf{H}^1(\tilde{\mathcal{J}}^{\mathrm{top}})$ is discrete, the inclusion homomorphism $\mathbf{H}^1(\tilde{\mathcal{J}}) \to \mathbf{H}^1(\tilde{\mathcal{J}}^{\mathrm{top}})$ is onto, and its kernel is connected. The last two statements follow from Corollary 4.1.5 and (an obvious extension of) the 5-lemma applied to the commutative diagram

$$\begin{tikzcd}
\mathbf{H}^0(\mathcal{S}) \arrow[r] \arrow[d, equal] & \mathbf{H}^1(\mathcal{J}) \arrow[r] \arrow[d, "i"', "\text{onto}"] & \mathbf{H}^1(\tilde{\mathcal{J}}) \arrow[r] \arrow[d, "\tilde\imath"'] & \mathbf{H}^1(\mathcal{S}) \arrow[r] \arrow[d, equal] & \mathbf{H}^2(\mathcal{J}) \arrow[d, "\cong"'] \\
\mathbf{H}^0(\mathcal{S}) \arrow[r] & \mathbf{H}^1(\mathcal{J}^{\mathrm{top}}) \arrow[r] & \mathbf{H}^1(\tilde{\mathcal{J}}^{\mathrm{top}}) \arrow[r] & \mathbf{H}^1(\mathcal{S}) \arrow[r] & \mathbf{H}^2(\mathcal{J}^{\mathrm{top}}).
\end{tikzcd}$$

One obtains that $\tilde\imath$ is an epimorphism and the induced homomorphism $\operatorname{Ker} i \to \operatorname{Ker}\tilde\imath$ is onto. The exact sequence in the statement of the theorem (and the discreteness of $\mathbf{H}^1(\tilde{\mathcal{J}}^{\mathrm{top}})$) follow then from the bottom row of the diagram via the identification $\mathbf{H}^i(\mathcal{J}^{\mathrm{top}}) = \mathbf{H}^{i+1}(R^1p_*\mathbb{Z}_J^-)$ given by Corollary 4.1.5. □

We conclude this subsection with an explanation of the count of real homological invariants corresponding to a fixed real functional invariant, see 3.2.6.

4.1.7 THEOREM. *If not empty, the set of real local systems belonging to and concordant with a given real map $j\colon B \to \mathbb{P}^1$ is an affine space over $\mathbf{H}^1(B^\sharp;\mathbb{Z}_2)$, where $B^\sharp = j^{-1}(\mathbb{P}^1 \smallsetminus \{0,1,\infty\})$.*

Proof. The proof repeats literally that of Theorem 4.1.1, with E and J replaced with, respectively, any and one of the local systems in question, and $\tilde{\mathcal{J}}$ replaced with the constant sheaf $(\mathbb{Z}_2)_{B^\sharp}$. □

Thus, the number of real local systems as in Theorem 4.1.7 is either 0 or 2^r, where $r = \dim \mathbf{H}^1(B^\sharp; \mathbb{Z}_2)$. The latter dimension is given by the following proposition.

4.1.8 Proposition. *For a connected complex curve B and for any real structure $c\colon B \to B$ one has*

$$\dim \mathbf{H}^1(B; \mathbb{Z}_2) = \dim (H^1(B; \mathbb{Z}_2))^c + 1 - \varepsilon,$$

where $\varepsilon = 1$ if $B_\mathbb{R} = \varnothing$ and either B is not compact or the genus of B is odd, and $\varepsilon = 0$ otherwise.

Proof. For $\mathbf{H}^1$, the spectral sequence (2.1.7) reduces to the exact sequence

$$0 \longrightarrow H^1(\mathbb{Z}_2; H^0(B; \mathbb{Z}_2)) \longrightarrow \mathbf{H}^1(B; \mathbb{Z}_2) \longrightarrow H^0(\mathbb{Z}_2; H^1(B; \mathbb{Z}_2)) \xrightarrow{\ {}^2d\ } \ldots.$$

Here, $H^1(\mathbb{Z}_2; H^0(B; \mathbb{Z}_2)) = \mathbb{Z}_2$ and $H^0(\mathbb{Z}_2; H^1(B; \mathbb{Z}_2)) = (H^1(B; \mathbb{Z}_2))^c$. Stabilizing (cf. 2.2.2), one observes that the differential

$${}^2d\colon H^0(\mathbb{Z}_2; H^1(B; \mathbb{Z}_2)) \to H^2(\mathbb{Z}_2; H^0(B; \mathbb{Z}_2)) = \mathbb{Z}_2$$

is nontrivial if and only if so is the differential ${}^2d\colon {}^2H^2 \to {}^2H^0$ in Kalinin's spectral sequence of B. If $B_\mathbb{R} \neq \varnothing$, then ${}^2d = 0$ as ${}^2H^0$ must survive to ${}^\infty H^0$. If $B_\mathbb{R} = \varnothing$ and B is not compact, i.e., $H^2(B; \mathbb{Z}_2) = 0$, then ${}^2d \neq 0$ as this is the only chance to kill ${}^2H^0 = \mathbb{Z}_2$. Finally, if $B_\mathbb{R} = \varnothing$ and B is compact, the involution is standard and a direct calculation shows that ${}^2d = 0$ if and only if the genus of B is even. □

4.2. The case of generic singular fibers. Fix a real Jacobian elliptic fibration $p\colon J \to B$ with a real section $s\colon B \to J$. Recall that, for any abelian group G, $R^0p_*G_J$ is the constant sheaf G_B and $R^2p_*G_J$ is an extension of G_B by a skyscraper sheaf $\mathcal{S}'$ concentrated at the points of B corresponding to reducible fibers of p.

4.2.1 Lemma. *For any abelian group G, the homomorphism $p^*\colon H^*(B; G) \to H^*(J; G)$* (equivalently, the edge homomorphism $H^*(B; R^0p_*G_J) \to H^*(J; G)$ of the Leray spectral sequence of p) *is a monomorphism and its image is a direct summand, the decomposition respecting c_J^*. Furthermore, p^* embeds Kalinin's spectral sequence of (B, c_B)* (both over $\mathbb{Z}_2$ and over $\mathbb{Z}$) *as a direct summand into Kalinin's spectral sequence of (J, c_J).*

Proof. All statements follow immediately from the existence of a section; the complementary direct summand is $\operatorname{Ker} s^*$, where s^* is the appropriate induced homomorphism. □

4.2.2 COROLLARY. *Starting from ${}^2H$, the only potentially nontrivial differential in the Leray spectral sequence of p is ${}^2d$: $H^0(B; R^2p_*G_J) \to H^2(B; R^1p_*G_J)$.* □

4.2.3 LEMMA. *If J is compact and nonisotrivial, then $H^0(B; R^1p_*\mathbb{Z}_J) = 0$ and for each $r \geq 1$, $m \in \mathbb{Z}_2$, and $q = 0, 1$ the induced homomorphism p^*: ${}^rH^{mq}(B; \mathbb{Z}) \to {}^rH^{mq}(J; \mathbb{Z})$ is an isomorphism. As a consequence, p^*: $\mathcal{F}^m_q(B; \mathbb{Z}) \to \mathcal{F}^m_q(J; \mathbb{Z})$ is an isomorphism for all $m \in \mathbb{Z}_2$ and $q = 0, 1$.*

Proof. The isomorphism p^*: ${}^1H^{m,0}(B; \mathbb{Z}) = {}^1H^{m,0}(J; \mathbb{Z})$ is obvious. Since J is nonisotrivial, its Mordel-Weil group is discrete, and from the exact sequence (3.2.9) one concludes that $H^0(B; R^1p_*\mathcal{O}_J) = 0$ and, hence, $H^0(B; R^1p_*\mathbb{Z}_J) = 0$. Thus, p^*: ${}^1H^{m,1}(B; \mathbb{Z}) \to {}^1H^{m,1}(J; \mathbb{Z})$ is also an isomorphism. Since p^* is a direct summand embedding, it remains an isomorphism for all $r \geq 1$. □

4.2.4 LEMMA. *If J is compact and nonisotrivial, then*

(1) *the groups $H_*(J; \mathbb{Z})$ and $H^*(J; \mathbb{Z})$ are torsion free;*

(2) *the "edge" homomorphism $H^3(J; \mathbb{Z}) \to H^1(B; R^2p_*\mathbb{Z}_J)$ is an isomorphism;*

(3) *the differential ${}^2d$: $H^0(B; R^2p_*\mathbb{Z}_J) \to H^2(B; R^1p_*\mathbb{Z}_J)$ establishes an isomorphism* $\operatorname{Coker}[H^2(J; \mathbb{Z}) \to H^0(B; R^2p_*\mathbb{Z}_J)] = H^2(B; R^1p_*\mathbb{Z}_J)$.

Proof. As is known (see, for example, [FM, Proposition 2.2.1]), the induced homomorphism p_*: $\pi_1(J) \to \pi_1(B)$ is an isomorphism. Hence, p_*: $H_1(J; \mathbb{Z}) \to H_1(B; \mathbb{Z})$ is also an isomorphism, the group $H_1(J; \mathbb{Z})$ is torsion free, and so are $H_*(J; \mathbb{Z})$ and $H^*(J; \mathbb{Z})$. Statement (2) follows then from the isomorphism p^*: $H^1(B; \mathbb{Z}) \to H^1(J; \mathbb{Z})$, see Lemma 4.2.3, and Poincaré duality, and (3) is immediate. □

Remark. An alternative description of the group $H^2(B; R^1p_*\mathbb{Z}_J)$, due to Kodaira, is as the group of coinvariants of p, i.e.,

$$H^2(B; R^1p_*\mathbb{Z}_J) = H^1(F_b; \mathbb{Z})/(1 - m_\gamma)a,$$

where F_b is the fiber over a point $b \in B^\sharp$, a runs through $H^1(F_b; \mathbb{Z})$, γ runs through $\pi_1(B^\sharp, b)$, and m_γ: $H^2(F_b; \mathbb{Z}) \to H^2(F_b; \mathbb{Z})$ is the monodromy along γ.

4.2.5 COROLLARY. *Assume that J is compact and nonisotrivial and that the real part $J_\mathbb{R}$ is nonempty* (equivalently, the real part $B_\mathbb{R}$ is nonempty). *Then J is both $\mathbb{Z}_2$- and $\mathbb{Z}$-Galois maximal.*

Proof. Since both $H^*(B; \mathbb{Z})$ and $H^*(J; \mathbb{Z})$ are torsion free, reduction modulo 2 induces isomorphisms of their Kalinin's spectral sequences with coefficients in $\mathbb{Z}$ and $\mathbb{Z}_2$, see 2.2.7. In particular, $\mathbb{Z}$- and $\mathbb{Z}_2$-Galois maximality are equivalent. Furthermore, Lemma 4.2.3 implies that p^*: ${}^rH^q(B; \mathbb{Z}_2) \to {}^rH^q(J; \mathbb{Z}_2)$ is an isomorphism for all $r \geq 1$ and $q = 0, 1$. The curve B with nonempty real part is $\mathbb{Z}_2$-Galois maximal. Hence, the differentials ${}^r d$, $r \geq 2$, landing in ${}^r H^q(J;\mathbb{Z}_2)$, $q = 0, 1$, are trivial, and using Poincaré duality 2.2.4 one concludes that so are all differentials ${}^r d$, $r \geq 2$, i.e., J is $\mathbb{Z}_2$-Galois maximal. □

4.2.6 LEMMA. *Assume that J is compact and nonisotrivial and that all its fibers are irreducible. Then*

(1) *one has $H^2(B; R^1 p_* \mathbb{Z}_J) = 0$;*

(2) *there are isomorphisms $\mathbf{H}^i(R^1 p_* \mathbb{Z}_J^-) = H^{i-1}(\mathbb{Z}_2; H^1(B; R^1 p_* \mathbb{Z}_J^-))$, $i \geq 0$,* (as usual $H^{-1} = 0$). *If, in addition, $B_{\mathbb{R}} \neq \varnothing$, then*

(3) *there is an isomorphism $\mathbf{H}^2(R^1 p_* \mathbb{Z}_J^-) = H^2(\mathbb{Z}_2; H^2(J;\mathbb{Z}))$.*

Proof. Pick a generic fiber F of p and let in: $F \hookrightarrow J$ be the inclusion. Denote by $f, b \in H_2(J;\mathbb{Z})$ the fundamental classes of F and $s(B)$, respectively. One has $f^2 = 0$ and $f \circ b = 1$. Hence, f and b span a unimodular sublattice in $H_2(J;\mathbb{Z})$ and there is an orthogonal (with respect to the intersection index form) direct sum decomposition $H_2(J;\mathbb{Z}) = \langle f, b\rangle \oplus \langle f, b\rangle^\perp$. The Poincaré duality yields then an orthogonal decomposition $H^2(J;\mathbb{Z}) = \operatorname{Hom}(\langle f, b\rangle, \mathbb{Z}) \oplus \operatorname{Ker}(\mathrm{in}^* \oplus s^*)$. In particular, the induced homomorphism $\mathrm{in}^*\colon H^2(J;\mathbb{Z}) \to H^2(F;\mathbb{Z})$ is onto.

Since all fibers are irreducible, one has $R^2 p_* \mathbb{Z}_J = \mathbb{Z}_B$ and $H^0(B; R^2 p_* \mathbb{Z}_J) = \mathbb{Z} = H^2(F;\mathbb{Z})$, and the edge homomorphism $H^2(J;\mathbb{Z}) \to H^0(B; R^2 p_* \mathbb{Z}_J)$ coincides with in^*. Since the latter is onto, 4.2.4(3) implies (1). Then $H^1(B; R^1 p_* \mathbb{Z}_J)$ is the only nontrivial cohomology of $R^1 p_* \mathbb{Z}_J$, and (2) follows from the spectral sequence (2.1.7). In particular, using the definition of $H^*(\mathbb{Z}_2; \cdot\,)$ given in 2.1.2, one has $\mathbf{H}^2(R^1 p_* \mathbb{Z}_J^-) = H^1(\mathbb{Z}_2; H^1(B; R^1 p_* \mathbb{Z}_J^-)) = H^2(\mathbb{Z}_2; H^1(B; R^1 p_* \mathbb{Z}_J))$.

Now, assume that $B_{\mathbb{R}} \neq \varnothing$. Then F can be chosen real, and the orthogonal decomposition above is c_J^*-equivariant. In view of Lemma 4.2.1, one can identify $H^1(B; R^1 p_* \mathbb{Z}_J)$ with $\operatorname{Ker}(\mathrm{in}^* \oplus s^*) \subset H^2(J;\mathbb{Z})$. Hence, one has

$$H^2(\mathbb{Z}_2; H^2(J;\mathbb{Z})) = H^2(\mathbb{Z}_2; \operatorname{Hom}(\langle f, b\rangle, \mathbb{Z})) \oplus H^2(\mathbb{Z}_2; H^1(B; R^1 p_* \mathbb{Z}_J)),$$

and it remains to notice that, F and $s(B)$ being analytic curves, c_* acts via minus identity on $\langle f, b\rangle$ and, hence, $H^2(\mathbb{Z}_2; \operatorname{Hom}(\langle f, b\rangle, \mathbb{Z})) = 0$. □

We are ready to prove the principal result of this section.

4.2.7 THEOREM. *Let $p\colon J \to B$ be a compact nonisotrivial real Jacobian elliptic surface with irreducible fibers and nonempty real part. Then there are canonical isomorphisms $\mathbb{R}\text{Ш}^{\mathrm{top}}(J) = H^2(\mathbb{Z}_2; H^2(J;\mathbb{Z})) = \mathcal{F}_2^0(J;\mathbb{Z})/\mathcal{F}_1^0(J;\mathbb{Z})$, where $\{\mathcal{F}_q^0\}$ is Kalinin's filtration, see 2.2.5.*

Alternatively, there is a canonical isomorphism $\mathbb{R}\text{Ш}^{\mathrm{top}}(J) = (\operatorname{Im} p^*)^\perp / \operatorname{Im} p^*$, where p^* is the induced homomorphism $H^{\mathrm{even}}(B_{\mathbb{R}};\mathbb{Z}_2) \to H^{\mathrm{even}}(J_{\mathbb{R}};\mathbb{Z}_2)$ and the complement is with respect to a certain perfect pairing $H^{\mathrm{even}} \otimes H^{\mathrm{even}} \to \mathbb{Z}_2$.

Proof. Under the assumptions one has $\mathcal{S} = 0$. Hence, Theorem 4.1.6 and Lemma 4.2.6(3) imply $\mathbb{R}\text{Ш}^{\mathrm{top}}(J) = \mathbf{H}^2(R^1 p_* \mathbb{Z}_J^-) = H^2(\mathbb{Z}_2; H^2(J;\mathbb{Z})) = {}^2H^{0,2}(J;\mathbb{Z})$. In view of Corollary 4.2.5, Kalinin's spectral sequence degenerates at ${}^2H$ and the latter group equals ${}^\infty H^{0,2}(J;\mathbb{Z}) = \mathcal{F}_2^0/\mathcal{F}_1^0$.

Table 1. The groups $H^i(\mathbb{Z}_2; H^1(F;\mathbb{Z}))$.

Fiber F	$H^1(\mathbb{Z}_2; H^1(F;\mathbb{Z}))$	$H^2(\mathbb{Z}_2; H^1(F;\mathbb{Z}))$
a nonsingular M-curve	$\mathbb{Z}_2$	$\mathbb{Z}_2$
a nonsingular $(M-1)$-curve	0	0
$\mathbb{P}^1/\{0\sim\infty\}$ with conj: $z\mapsto\bar z$	$\mathbb{Z}_2$	0
$\mathbb{P}^1/\{0\sim\infty\}$ with conj: $z\mapsto 1/\bar z$	0	$\mathbb{Z}_2$
rational curve with a cusp	0	0

Since the group $H^*(J;\mathbb{Z})$ is torsion free, see 4.2.4(1), the map $1+\mathrm{Sq}^1$ establishes isomorphisms $\mathcal{F}^0_q(J;\mathbb{Z}) = (1+\mathrm{Sq}^1)\mathcal{F}_q(J;\mathbb{Z}_2)\cap H^{\mathrm{even}}(J_{\mathbb{R}};\mathbb{Z}_2)$, see 2.2.7. On the other hand, with respect to the pairing described in 2.2.4, one has $\mathcal{F}_2 = \mathcal{F}_1^{\perp}$, and, in view of Lemma 4.2.3, the map $p^*\colon\ \mathcal{F}_1(B;\mathbb{Z}_2)\to\mathcal{F}_1(J;\mathbb{Z}_2)$ is an isomorphism. Finally, since B is a compact curve, the intersection $(1+\mathrm{Sq}^1)\mathcal{F}_1(B;\mathbb{Z}_2)\cap H^{\mathrm{even}}(B_{\mathbb{R}};\mathbb{Z}_2)$ is merely the group $H^0(B_{\mathbb{R}};\mathbb{Z}_2) = H^{\mathrm{even}}(B_{\mathbb{R}};\mathbb{Z}_2)$. $\square$

4.3. The geometric interpretation. Numerically, Theorem 4.2.7 states that, under the assumptions, the discrete part $\mathbb{R}\text{Ш}^{\mathrm{top}}(J)$ is a $\mathbb{Z}_2$-vector space of dimension twice the number of extra components of $J_{\mathbb{R}}$. Indeed, each component of $J_{\mathbb{R}}$ contributes 2 to $\dim H^{\mathrm{even}}(J_{\mathbb{R}};\mathbb{Z}_2)$, and the passage to $(\operatorname{Im} p^*)^{\perp}/\operatorname{Im} p^*$ kills the contribution of the principal components. Below we show that, in fact, each extra component of $J_{\mathbb{R}}$ does contribute a pair of $\mathbb{Z}_2$ summands in a natural way.

As in Theorem 4.2.7, fix a compact nonisotrivial real Jacobian elliptic surface $p\colon J\to B$ with irreducible fibers and nonempty real part. Let $J' = p^{-1}(B_{\mathbb{R}})$, and regard $p\colon J'\to B_{\mathbb{R}}$ as a real "Jacobian fibration." In particular, one can consider the discrete Tate-Shafarevich group $\mathbb{R}\text{Ш}^{\mathrm{top}}(J')$. Clearly, Theorem 4.1.6 still applies and yields a natural isomorphism $\mathbb{R}\text{Ш}^{\mathrm{top}}(J') = \mathbf{H}^2(R^1p_*\mathbb{Z}^-_{J'})$. Note that the sheaf $R^1p_*\mathbb{Z}^-_{J'}$ is the restriction $i^*R^1p_*\mathbb{Z}^-_J$, where $i\colon B_{\mathbb{R}}\to B$ is the inclusion.

4.3.1 Lemma. *Let J and J' be as above. Then the inclusion homomorphism $i^*\colon\ \mathbf{H}^2(R^1p_*\mathbb{Z}^-_J)\to\mathbf{H}^2(R^1p_*\mathbb{Z}^-_{J'})$ is onto.*

Proof. Let $\mathcal{K}$ be the kernel of the epimorphism $R^1p_*\mathbb{Z}^-_J\to i_*R^1p_*\mathbb{Z}^-_{J'}$. It suffices to show that $\mathbf{H}^3(\mathcal{K})=0$. Since $\mathcal{K}$ is trivial over $B_{\mathbb{R}}$, one has $\mathcal{H}^p(\mathbb{Z}_2;\mathcal{K})=0$ for all $p>0$, and the spectral sequence (2.1.6) reduces to the isomorphisms $\mathbf{H}^q(\mathcal{K}) = H^q(B/c;\mathcal{H}^0(\mathbb{Z}_2;\mathcal{K}))$. As $3>\dim B/c = 2$, the statement follows. $\square$

Similarly, one can speak about the discrete Tate-Shafarevich group $\mathbb{R}\text{Ш}^{\mathrm{top}}(F)$ of a single fiber $F=p^{-1}(b)$, $b\in B_{\mathbb{R}}$. As above, for an irreducible fiber F one has $\mathbb{R}\text{Ш}^{\mathrm{top}}(F) = \mathbf{H}^2(R^1p_*\mathbb{Z}^-_F) = H^1(\mathbb{Z}_2;H^1(F;\mathbb{Z}))$. (As usual, the dimension shift is due to the nonstandard action of the involution on $\mathbb{Z}^-_F$.) This group is either $\mathbb{Z}_2$ or 0, see Table 1. In the case $\mathbb{R}\text{Ш}^{\mathrm{top}}(F)=\mathbb{Z}_2$, the nontrivial element corresponds to the real structure on F with respect to which the normalization of F has empty real part.

Denote by $J^{\mathrm{ext}}_{\mathbb{R}}$ the union of the extra components of $J_{\mathbb{R}}$, and let $B^{\mathrm{ext}}_{\mathbb{R}} = p_*(J^{\mathrm{ext}}_{\mathbb{R}})$.

4.3.2 THEOREM. *Let $p\colon J \to B$ be a compact nonisotrivial real Jacobian elliptic surface with irreducible fibers and nonempty real part, and let $J' = p^{-1}(B_{\mathbb{R}})$. Then there is a natural exact sequence*

$$\begin{aligned} 0 \longrightarrow H^1(B_{\mathbb{R}}; \mathcal{H}^1(\mathbb{Z}_2; R^1p_*\mathbb{Z}^-_{J'})) &\xrightarrow{\alpha} \mathbb{R}\text{Ш}^{\mathrm{top}}(J') \xrightarrow{\beta} \\ &\xrightarrow{\beta} H^0(B_{\mathbb{R}}; \mathcal{H}^2(\mathbb{Z}_2; R^1p_*\mathbb{Z}^-_{J'})) \longrightarrow 0 \end{aligned}$$

and natural isomorphisms

$$\begin{aligned} H^1(B_{\mathbb{R}}; \mathcal{H}^1(\mathbb{Z}_2; R^1p_*\mathbb{Z}^-_{J'})) &= H^1(B^{\mathrm{ext}}_{\mathbb{R}}, \partial B^{\mathrm{ext}}_{\mathbb{R}}; \mathbb{Z}_2), \\ H^0(B_{\mathbb{R}}; \mathcal{H}^2(\mathbb{Z}_2; R^1p_*\mathbb{Z}^-_{J'})) &= H^0(B^{\mathrm{ext}}_{\mathbb{R}}; \mathbb{Z}_2) = \bigoplus \mathbb{R}\text{Ш}^{\mathrm{top}}(F_i), \end{aligned}$$

where $F_i = p^{-1}(b_i)$ are the fibers over some points $b_i \in B_{\mathbb{R}}$, one in the interior of each connected component of $B^{\mathrm{ext}}_{\mathbb{R}}$. The composition of the last isomorphism and β coincides with the homomorphism induced by the inclusion $\bigcup\{b_i\} \hookrightarrow B_{\mathbb{R}}$.

Proof. Since $\dim B_{\mathbb{R}} = 1$, one has $H^q(B_{\mathbb{R}}; \cdot\,) = 0$ for $q > 1$ and, hence, the spectral sequence (2.1.6) for $\mathbb{R}\text{Ш}^{\mathrm{top}}(J') = \mathbf{H}^2(R^1p_*\mathbb{Z}^-_{J'})$ collapses at E_2 and results in the exact sequence in the statement.

The stalks of the sheaves $\mathcal{H}^p(\mathbb{Z}_2; R^1p_*\mathbb{Z}^-_{J'})$, $p = 1, 2$, are given by Table 1. The stalks are at most $\mathbb{Z}_2$, and the sheaves are supported by the closure $\bar{B}^M_{\mathbb{R}}$ of the subset $B^M_{\mathbb{R}}$ of the points of $B_{\mathbb{R}}$ whose pull-back is a nonsingular M-curve. Hence, the cohomology groups are $H^q(B_{\mathbb{R}}; \cdot\,) = H^q(\bar{B}^M_{\mathbb{R}}, D; \mathbb{Z}_2)$, where D is the part of the boundary $\partial\bar{B}^M_{\mathbb{R}}$ over which the sheaf in question has trivial stalks. Considering the possibilities for a component of $\bar{B}^M_{\mathbb{R}}$, one easily concludes that only the components of $B^{\mathrm{ext}}_{\mathbb{R}} \subset \bar{B}^M_{\mathbb{R}}$ make nontrivial contributions to the cohomology; this observation gives the isomorphisms in the statement. □

4.3.3 COROLLARY. *Under the hypothesis of Theorem 4.3.2, the inclusion homomorphism $i^*\colon \mathbb{R}\text{Ш}^{\mathrm{top}}(J) \to \mathbb{R}\text{Ш}^{\mathrm{top}}(J')$ is an isomorphism.*

Proof. According to Theorem 4.3.2, each extra component of $J_{\mathbb{R}}$ contributes 4 to the order of $\mathbb{R}\text{Ш}^{\mathrm{top}}(J')$. Hence, the two groups are of the same order, and the statement follows from Lemma 4.3.1. □

4.3.4. Combining Theorem 4.3.2 and the proof of Theorem 4.1.1 (namely, the part explaining how a real elliptic fibration can be modified), one can easily describe the real parts of all, not necessarily Jacobian, elliptic surfaces. Each extra component $X \subset J_{\mathbb{R}}$ contributes two $\mathbb{Z}_2$ summands to $\mathbb{R}\text{Ш}(J)$: one to the subgroup $H^1 = H^1(B_{\mathbb{R}}; \mathcal{H}^1(\mathbb{Z}_2; R^1p_*\mathbb{Z}^-_{J'}))$, and one to the quotient $H^0 = H^0(B_{\mathbb{R}}; \mathcal{H}^2(\mathbb{Z}_2; R^1p_*\mathbb{Z}^-_{J'}))$. The nontrivial element of H^1 represents a real modification of the fibration, given by a real 1-cocycle, cf. the proof of Theorem 4.1.1. If X is a sphere, Figure 2(a), the resulting modification of the real part is shown

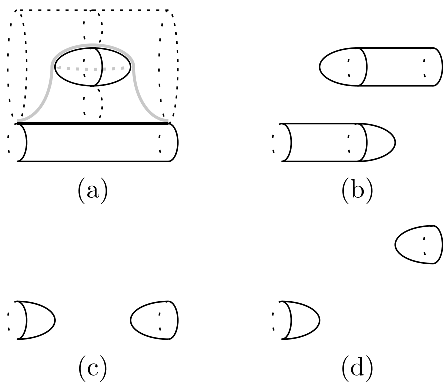


Figure 2. Modifications of the real part.

in Figure 2(b), the cocycle being the partial section shown by a gray dotted line in Figure 2(a). If X is a torus or a Klein bottle, the two components over the corresponding hyperbolic component of $B_{\mathbb{R}}$ are intertwined into one. Note that, since the restriction of H^1 to each fiber is trivial, the real structures of all fibers remain intact.

The nontrivial element of H^0 contributed by X restricts nontrivially to each fiber F over the projection $p(X) \subset B_{\mathbb{R}}$, thus resulting in rebuilding the real structure of all fibers. If X is a torus or a Klein bottle, the result has empty real part (over the corresponding hyperbolic component of $B_{\mathbb{R}}$). If X is a sphere, the result is shown in Figure 2(c) and (d). The new real structure is obtained via the shift by a real section of the opposite Jacobian fibration. One such section, which is homotopically nontrivial in each fiber over $p(X)$, is shown in gray in Figure 2(a). Note that the real parts shown in Figure 2(c) and (d) are homeomorphic; the two figures are intended to indicate the fact that one can just change the real structure or change both the real structure and the fibration.

4.3.5 Corollary. *A compact nonisotrivial real elliptic surface $p\colon E \to B$ with irreducible fibers is equivariantly deformation equivalent to a real Jacobian surface if and only if the restriction $p\colon E_{\mathbb{R}} \to B_{\mathbb{R}}$ admits a continuous section.*

Proof. The "only if" part is obvious. For the "if" part, it suffices to notice that none of the fibrations shown in Figure 2(b)–(d) (neither the results of modifying a toroidal extra component) has a section. □

4.3.6 Proposition. *Let E be a compact nonisotrivial real elliptic surface that is not equivariantly deformation equivalent to its Jacobian $J = J(E)$. Then*

$$\dim H^*(E_{\mathbb{R}}; \mathbb{Z}_2) \leq \dim H^*(J_{\mathbb{R}}; \mathbb{Z}_2) - 4.$$

Proof. Each modification shown on Figure 2, from (a) to any of (b), (c), (d), either leaves the total Betti number $\dim H^*(\,\cdot\,; \mathbb{Z}_2)$ unchanged or reduces it by 4, depending on whether a new component is created or not. As the first modification of this kind, starting from $J = J(E)$, does not create a new component, it does reduce the total Betti number. Similarly, each nontrivial modification over a hyperbolic component of $B_{\mathbb{R}}$ turns a pair of tori (or Klein bottles) into a single torus (respectively, Klein bottle) or the empty set, thus reducing the total Betti number by 4 or 8, respectively. □

4.3.7 COROLLARY. *A compact nonisotrivial real elliptic surface that is an M- or $(M-1)$-variety is equivariantly deformation equivalent to its Jacobian.* □

Remark. The same arguments as above show that under the same hypotheses the inequality

$$\dim H^k(E_{\mathbb{R}};\mathbb{Z}_2) \le \dim H^k(J_{\mathbb{R}};\mathbb{Z}_2),$$

where $J = J(E)$ is the Jacobian, holds for any k.

4.4. Deformations. Let $p\colon X \to S$, $\pi\colon S \to D$ be an elementary real deformation of nonisotrivial compact real Jacobian elliptic surfaces. The projection $p\colon X \to S$ can be regarded as an elliptic fibration; hence, one can define the corresponding real Tate-Shafarevich group $\mathbb{R}Ш(X)$, the sheaves $\tilde{\mathcal{J}}$ and $\mathcal{J} = R^1p_*\mathcal{O}_X/R^1p_*\mathbb{Z}_X^-$, and their topological counterparts $\mathbb{R}Ш(X)^{\mathrm{top}}$, $\tilde{\mathcal{J}}^{\mathrm{top}}$, and $\mathcal{J}^{\mathrm{top}}$. As in Theorems 4.1.1 and 4.1.6, one has $\mathbb{R}Ш(X) = \mathbf{H}^1(\tilde{\mathcal{J}})$ and $\mathbb{R}Ш^{\mathrm{top}}(X) = \mathbf{H}^1(\tilde{\mathcal{J}}^{\mathrm{top}})$.

4.4.1 THEOREM. *Assume that the members $p_t\colon X_t \to S_t$, $t \in D$, of the family, except possibly X_0, have no singular fibers other than those of type* I_1. *Then the restriction homomorphism $\mathbb{R}Ш(X) \to \mathbb{R}Ш(X_0)$ is onto, and the restriction homomorphism $\mathbb{R}Ш^{\mathrm{top}}(X) \to \mathbb{R}Ш^{\mathrm{top}}(X_0)$ is an isomorphism.*

4.4.2 COROLLARY. *Let E be a compact real elliptic surface, J its Jacobian, and $p\colon X \to D$, $\pi\colon S \to D$ a real deformation of J* (i.e., $X_0 = J$) *as in Theorem 4.4.1. Then there is a real deformation $\tilde{p}\colon \tilde{X} \to D$ of E* (with the same base $\pi\colon S \to D$) *such that $X_t = J(\tilde{X}_t)$ for each $t \in D$.* □

Proof of Theorem 4.4.1. Consider the sheaves $\mathcal{J}$, $\tilde{\mathcal{J}}$, and $\mathcal{S}$ on S and denote by $\mathcal{J}_0$, $\tilde{\mathcal{J}}_0$, and $\mathcal{S}_0$, respectively, their restrictions to S_0. There are exact sequences

$$0 \longrightarrow \mathcal{J} \longrightarrow \tilde{\mathcal{J}} \longrightarrow \mathcal{S} \longrightarrow 0, \qquad 0 \longrightarrow \mathcal{J}_0 \longrightarrow \tilde{\mathcal{J}}_0 \longrightarrow \mathcal{S}_0 \longrightarrow 0,$$

and, since the restriction homomorphisms $\mathbf{H}^*(\mathcal{S}) \to \mathbf{H}^*(\mathcal{S}_0)$ are isomorphisms (as $\mathcal{S}$ is concentrated at points of S_0), it suffices to show that the homomorphism $\mathbf{H}^i(\mathcal{J}) \to \mathbf{H}^i(\mathcal{J}_0)$ is onto for $i = 1$ and one-to-one for $i = 2$.

To this end, we compare the exact sequences

$$0 \longrightarrow R^1 p_* \mathbb{Z}_X^- \longrightarrow R^1 p_* \mathcal{O}_X \longrightarrow \mathcal{J} \longrightarrow 0,$$

$$0 \longrightarrow R^1 p_* \mathbb{Z}_{X_0}^- \longrightarrow R^1 p_* \mathcal{O}_{X_0} \longrightarrow \mathcal{J}_0 \longrightarrow 0.$$

For the coherent sheaves $R^1 p_* \mathcal{O}_X$ and $R^1 p_* \mathcal{O}_{X_0}$ one has $\mathbf{H}^i(\,\cdot\,) = H^0(\mathbb{Z}_2; H^i(\,\cdot\,))$ and $H^j(\mathbb{Z}_2; H^i(\,\cdot\,)) = 0$ for $j > 0$ (cf. Lemma 4.1.4). Furthermore, for each fiber S_t of π one has $H^j(S_t; R^1 p_* \mathcal{O}_{X_t}) = 0$ unless $j = 1$ (see 3.2.7), and the Leray spectral sequence of π implies that $H^i(X; R^1 p_* \mathcal{O}_X) = H^{i-1}(D; R^1 \pi_* R^1 p_* \mathcal{O}_X)$. Since D is a Stein manifold, one concludes that $\mathbf{H}^2(R^1 p_* \mathcal{O}_X) = \mathbf{H}^2(R^1 p_* \mathcal{O}_{X_0}) = 0$ and the homomorphism $\mathbf{H}^1(R^1 p_* \mathcal{O}_X) \to \mathbf{H}^1(R^1 p_* \mathcal{O}_{X_0})$ is onto. Thus, it remains to prove that the restriction homomorphism $\mathbf{H}^1(R^1 p_* \mathbb{Z}_X^-) \to \mathbf{H}^1(R^1 p_* \mathbb{Z}_{X_0}^-)$ is also onto. We will show that it is, in fact, an isomorphism.

Informally, the last assertion follows from the isomorphism $H^*(X) = H^*(X_0)$. More precisely, observe that $R^0 p_* \mathbb{Z}_X$ and $R^0 p_* \mathbb{Z}_{X_0}$ are both constant sheaves (with fiber $\mathbb{Z}$) and $R^2 p_* \mathbb{Z}_X$ and $R^2 p_* \mathbb{Z}_E$ are extensions of constant sheaves by a sheaf concentrated at points of X_0. Hence, both $H^*(R^0 p_* \mathbb{Z}_X) \to H^*(R^0 p_* \mathbb{Z}_{X_0})$ and $H^*(R^2 p_* \mathbb{Z}_X) \to H^*(R^2 p_* \mathbb{Z}_{X_0})$ are isomorphisms. From comparing the Leray spectral sequences for the projections $X \to S$ and $X_0 \to S_0$ it follows that the maps $H^i(R^1 p_* \mathbb{Z}_X) \to H^i(R^1 p_* \mathbb{Z}_E)$ are also isomorphisms (for all i), and the statement follows from the spectral sequence (2.1.7).

The proof of the topological statement is similar, except that for fine sheaves one has $H^j(S_t; (R^1 p_* \mathcal{O}_{X_t})^{\mathrm{top}}) = 0$ unless $j = 0$ (for each $t \in D$) and, hence, $\mathbf{H}^i(R^1 p_* \mathcal{O}_X) = \mathbf{H}^i(R^1 p_* \mathcal{O}_{X_0}) = 0$ for $i = 1$, 2. Alternatively, the assertion can be derived from Theorem 4.1.6 and a similar exact sequence for $\mathbb{R}\text{Ш}^{\mathrm{top}}(X)$. □

5. Real trigonal curves and dessins d'enfants. We start this section by introducing the notion of trichotomic graph. It is a real version of Grothendieck's *dessins d'enfants*, which is adjusted for dealing with real meromorphic functions defined on a real curve and having a certain preset ramification over the three real points $0, 1, \infty \in \mathbb{P}^1$. More precisely, a trichotomic graph is the quotient by the complex conjugation of a properly decorated pull-back of $(\mathbb{P}^1_{\mathbb{R}}; 0, 1, \infty)$; the pull-backs of 0, 1, and ∞ being marked with $\bullet$-, $\circ$-, and $\times$- respectively. Note that the function may (and usually does) have other ramification points, which are ignored unless they are real.

5.1. Trichotomic graphs. Let D be a (topological) compact connected surface, possibly with boundary. (Unless specified otherwise, in the topological part of this section we are working in the ***PL***-category.) We use the term *real* for points, segments, etc. situated at the boundary ∂D. For a graph $\Gamma \subset D$, we denote by D_Γ the closed cut of D along Γ. The connected components of D_Γ are called *regions* of Γ.

A *trichotomic graph* on D is an embedded oriented graph $\Gamma \subset D$ decorated with the following additional structures (referred to as *colorings* of the edges and vertices of Γ, respectively):

• each edge of Γ is of one of the three kinds: solid, bold, or dotted;

• each vertex of Γ is of one of the four kinds: $\bullet$, $\circ$, $\times$, or monochrome (the vertices of the first three kinds being called *essential*);

and satisfying the following conditions:

(1) the boundary ∂D is a union of edges and vertices of Γ;

(2) the valency of each essential vertex of Γ is at least 2, and the valency of each monochrome vertex of Γ is at least 3;

(3) the orientations of the edges of Γ form an orientation of the boundary ∂D_Γ; this orientation extends to an orientation of D_Γ;

(4) all edges incident to a monochrome vertex are of the same kind;

(5) $\times$-vertices are incident to incoming dotted edges and outgoing solid edges;

(6) $\bullet$-vertices are incident to incoming solid edges and outgoing bold edges;

(7) $\circ$-vertices are incident to incoming bold edges and outgoing dotted edges.

(8) each triangle (i.e., region with three essential vertices in the boundary) is a topological disk.

In (5)–(7) the lists are complete, i.e., vertices cannot be incident to edges of other kinds or with different orientation.

In view of (4), the monochrome vertices can further be subdivided into solid, bold, and dotted, according to their incident edges. The sets of solid, bold, and dotted monochrome vertices of Γ will be denoted by Γ_{solid}, Γ_{bold}, and Γ_{dotted}, respectively. The *monochrome part* of Γ of a given kind (solid, bold, or dotted) is the union of (open) edges and monochrome vertices of the corresponding kind. Thus, essential vertices **never** belong to a monochrome part.

Condition (3) implies, in particular, that the orientations of the edges incident to a vertex alternate. (This statement is equivalent to the first part of (3).) Thus, all inner vertices of Γ have even valencies.

A number of examples of trichotomic graphs is found further in this section (see, e.g., Figure 26, where complete graphs are drawn).

5.1.1. Let Γ be a trichotomic graph on D. If D is orientable, a choice of the orientation defines a chessboard coloring of D_Γ: a region $D_i \subset D_\Gamma$ is said to be *positive* (*negative*) if its orientation induced from D coincides with (respectively, is opposite to) that defined by Γ. Conversely, a chessboard coloring of D_Γ defines an orientation of D.

5.1.2. A path in a trichotomic graph Γ is called *monochrome* if it belongs to a monochrome part of Γ. Given two monochrome vertices $u, v \in \Gamma$, we say that $u \prec v$ if there is an oriented monochrome path from u to v. (Clearly, only vertices of the same kind can be compatible.) The graph is called *admissible* if $\prec$ is a partial order. Since $\prec$ is obviously transitive, this condition is equivalent to the requirement that Γ should have no oriented monochrome cycles.

Remark. Note that the orientation of Γ is almost superfluous. Indeed, Γ may have at most two orientations satisfying (3), and if Γ has at least one essential vertex, its orientation is uniquely determined by (5)–(7). Note also that (each connected component of) an admissible graph does have essential vertices, as otherwise any component of ∂D_Γ would be an oriented monochrome cycle.

Remark. In fact, all three decorations of an admissible graph Γ (orientation and the two colorings) can be recovered from any of the colorings. However, for clarity we retain both colorings in the diagrams.

5.1.3. Let B be the orientable double of D, i.e., the orientation double covering of D with the two preimages of each real point $d \in \partial D$ identified. (We exclude the case when D is closed and orientable, as then B would be disconnected.) Denote by $p\colon B \to D$ the projection and by $c\colon B \to B$ its deck translation, which is an orientation reversing involution. We will show that trichotomic graphs on D are merely a way of describing c-invariant trichotomic graphs on B. For this purpose, given a trichotomic graph Γ on D, consider its pull-back $\Gamma' = p^{-1}(\Gamma)$ and equip it with the decorations induced by p. Clearly, the deck translation c **preserves** the decorations of Γ', including its orientation.

5.1.4 LEMMA. *Given a trichotomic graph $\Gamma \subset D$, its pull-back $\Gamma' = p^{-1}(\Gamma)$, with the decorations induced by p, is a c-invariant trichotomic graph on B. Conversely, given a c-invariant trichotomic graph $\Gamma' \subset B$, its quotient $\Gamma = p(\Gamma')$ is a trichotomic graph on D. The graph Γ' is admissible if and only if so is Γ.*

Proof. The direct statement is immediate. For the converse, assume that $\Gamma' \subset B$ is a c-invariant trichotomic graph. Since c is orientation reversing, the graph Γ' has the following **separation property**: each region B_i of Γ' is disjoint from its image $c(B_i)$. (In fact, B_i and $c(B_i)$ have opposite signs in the sense of 5.1.1.) Hence, the restriction of p to B_i is a one-to-one map onto a region $D_i \subset D_\Gamma$. Since, in addition, the restriction $p\colon \Gamma' \to \Gamma$ is orientation preserving, property (3) for Γ follows from (3) for Γ'. The separation property implies also that Γ' contains the fixed point set $\operatorname{Fix} c$; this yields (1) and (2) for Γ.

Since p preserves the decorations of Γ' and Γ, the admissibility of one of the graphs implies the admissibility of the other. □

The *full valency* of a vertex of Γ is the valency of any of its pull-backs in Γ'. The full valency of an inner vertex coincides with its valency; the full valency of a real vertex equals $2 \cdot \text{valency} - 2$. The full valency of any vertex is even.

In what follows, we denote by $\#_\circ(\Gamma)$, $\#_\bullet(\Gamma)$, and $\#_\times(\Gamma)$ the numbers of, respectively, $\circ$-, $\bullet$-, and $\times$-vertices of Γ'. These numbers can be regarded as weighted numbers of respective vertices in Γ, each inner vertex being counted twice.

5.1.5. A typical example of a trichotomic graph is the following. Let B be a connected closed surface with involution c, and let $j\colon (B,c) \to (\mathbb{P}^1, {}^-)$ be an equivariant ramified covering. (Thus, B is necessarily orientable, c is orientation reversing, and one can assume j orientation preserving.) Then j defines a trichotomic graph $\Gamma'(j) \subset B$. As a set, $\Gamma'(j)$ is the pull-back $j^{-1}(\mathbb{P}^1_{\mathbb{R}})$. The trichotomic graph structure on $\Gamma'(j)$ is introduced as follows: the $\bullet$-, $\circ$-, and $\times$-vertices are the pull-backs of 0, 1, and ∞, respectively (monochrome vertices being the branch points with other real critical values), the edges are solid, bold, or dotted provided that their images belong to $[\infty, 0]$, $[0, 1]$, or $[1, \infty]$, respectively, and the orientation of $\Gamma'(j)$ is that induced from the positive orientation of $\mathbb{P}^1_{\mathbb{R}}$ (i.e., order of $\mathbb{R}$).

5.1.6 LEMMA. *The graph $\Gamma'(j) \subset B$ constructed above is an admissible c-invariant trichotomic graph. Hence, its image $\Gamma(j) = \Gamma'(j)/c \subset B/c = D$ is an admissible trichotomic graph.*

Proof. By Lemma 5.1.4, the second statement follows from the first one. Axiom (3) for $\Gamma'(j)$ follows from the fact that a region $B_i \subset B_{\Gamma'(j)}$ is positive (negative) in the sense of 5.1.1 if its image is the disk $\{\operatorname{Im} z \geq 0\}$ (respectively, $\{\operatorname{Im} z \leq 0\}$). Other axioms are straightforward. The admissibility follows from the fact that, since $j\colon \Gamma'(j) \to \mathbb{P}^1_{\mathbb{R}}$ is orientation preserving, $\prec$ is a subset of the partial order induced by the linear orders on the intervals $(1, \infty)$, $(\infty, 0)$, and $(0, 1)$. □

5.1.7 THEOREM. *Let D be a compact connected surface and (B, c) its orientable double. Exclude the case of oriented D without boundary, and equip B with its canonical orientation. Then a trichotomic graph $\Gamma \subset D$ is admissible if and only if it has the form $\Gamma(j)$ for some orientation preserving equivariant ramified covering $j\colon (B, c) \to (\mathbb{P}^1, {}^-)$. Furthermore, j is determined by Γ up to homotopy in the class of equivariant ramified covering having a fixed trichotomic graph.*

Proof. The "if" part is given by Lemma 5.1.6. For the "only if" part, we will construct a map j and, at each step, check that the construction is unique up to homotopy.

Any map j in question must have an orientation preserving descent $\tilde{\jmath}\colon \Gamma \to \mathbb{P}^1_{\mathbb{R}}$. The images of the essential vertices are predefined, and the extension of $\tilde{\jmath}$ to, say, the solid part of the graph is determined, up to homotopy, by a monotonous map $\tilde{\jmath}\colon (\Gamma_{\text{solid}}, \prec) \to ((1, \infty), <)$. The set of such maps is defined by linear inequalities $\tilde{\jmath}(u) < \tilde{\jmath}(v)$ whenever $u \prec v$, $u, v \in \Gamma_{\text{solid}}$. Hence, as a convex subset of a Cartesian power of $(1, \infty)$, it is connected. For the existence, one can, e.g., extend $\prec$ to a linear order (any maximal order) on Γ_{solid} and map the vertices to consecutive integers.

Let $\Gamma' \subset B$ be the pull-back of Γ, see 5.1.3 for the notation. The composition $\tilde{\jmath} \circ p$ is an equivariant orientation preserving map $j\colon (\Gamma', c) \to (\mathbb{P}^1_{\mathbb{R}}, {}^-)$. For each positive (in the sense of 5.1.1) region B_i of Γ', the restriction $j\colon \partial B_i \to \mathbb{P}^1_{\mathbb{R}}$ is a covering; since the orientations on B_i and ∂B_i agree, j extends to an orientation preserving ramified covering $j\colon B_i \to \{\operatorname{Im} z \geq 0\}$. Then $^{-} \circ j \circ c\colon c(B_i) \to \{\operatorname{Im} z \leq 0\}$ extends j to the negative components. The separation property of Γ' (see 5.1.4) assures that the extension $j\colon B \to \mathbb{P}^1$ is well defined and equivariant. Each inner point of an edge of Γ' is regular (as adjacent components of $B_{\Gamma'}$ have opposite signs); hence, j has isolated critical points and thus is a ramified covering.

The only ambiguity in the last step of the construction is in extending a covering $\partial B_i \to \mathbb{P}^1_{\mathbb{R}}$ of the circle to a ramified covering $B_i \to \{\operatorname{Im} z \geq 0\}$ of the disk. Any such extension can be perturbed to a generic one (with all branch points double and all critical values distinct), and the latter is unique up to homotopy due to an analog of the Hurwitz theorem (see, e.g., [P] or [N2]; a very transparent proof is indicated in [BE]). □

Theorem 5.1.7 and the Riemann existence theorem result in the following corollary.

5.1.8 COROLLARY. *Given an admissible trichotomic graph $\Gamma \subset D$, there is a complex structure on B and a holomorphic map $j\colon B \to \mathbb{P}^1$ such that the canonical orientation of B coincides with its complex orientation, c is a real structure on B, $j\colon (B, c) \to (\mathbb{P}^1, ^{-})$ is equivariant, and $\Gamma = \Gamma(j)$. Both the complex structure and the map are unique up to deformation.* □

Remark. As it follows from the proof, a slightly stronger statement holds. On each of the sets Γ_{solid}, Γ_{bold}, Γ_{dotted} one can fix in advance a partial order extending $\prec$. Then j can be chosen compatible with the given partial orders, and j is unique up to homotopy in the class of such maps.

5.2. Deformations. Let us fix an oriented closed connected surface B with an orientation reversing involution $c\colon B \to B$. Let $D = B/c$ and let $p\colon B \to D$ be the projection. We are interested in orientation preserving equivariant ramified coverings $j\colon (B, c) \to (\mathbb{P}^1, ^{-})$. A *deformation* of coverings is a homotopy $B \times I \to \mathbb{P}^1$ in the class of equivariant ramified coverings. A deformation is called *simple* if it preserves the multiplicities of all the points with values 0, 1, and ∞ and the multiplicities of all branch points with real critical values. Clearly, any deformation is locally simple with the exception of finitely many isolated values of the parameter $t \in I$. (As in Section 5.1, we are working in the ***PL***-category; in particular, this implies the finiteness.) The following statement is an immediate consequence of Theorem 5.1.7 and the definition of $\Gamma(j)$.

5.2.1 PROPOSITION. *Two equivariant ramified coverings $j_0, j_1\colon B \to \mathbb{P}^1$ can be connected by a simple deformation if and only if their graphs $\Gamma(j_0)$, $\Gamma(j_1)$ are isotopic.* □

Let $\Gamma_0 \subset D$ be a trichotomic graph. Pick some disjoint regular neighborhoods U_v of all (or some) vertices v of Γ_0 (we assume that $U_v \cap \partial D = \varnothing$ unless

v is real) and replace each intersection $\Gamma_0 \cap U_v$ with another decorated graph, so that the result Γ_1 is again a trichotomic graph. If each intersection $\Gamma_1 \cap U_v$ contains essential vertices of at most one kind, Γ_1 is called a *perturbation* of Γ_0 (and Γ_0 is called a *degeneration* of Γ_1). A perturbation Γ_1 of an admissible trichotomic graph Γ_0 is admissible if and only if none of the intersections $\Gamma_1 \cap U_v$ contains an oriented monochrome cycle. (Note that there are no simple local criteria for the admissibility of a degeneration.)

Remark. Assume that Γ_1 is a perturbation of Γ_0, and $\Gamma_1 \cap U_v$ contains no oriented monochrome cycles. Since the intersection $\Gamma_1 \cap \partial U_v$ is fixed, the assumption on $\Gamma_1 \cap U_v$ implies that $\Gamma_1 \cap U_v$ either is monochrome (if v is monochrome) or consists of monochrome vertices, essential vertices of the same kind as v, and edges of the two kinds incident to v.

Any deformation j_t of ramified coverings whose restriction to $B \times (0, 1]$ is simple results in a perturbation of the graph $\Gamma_0 = \Gamma(j_0)$. (The requirement that each intersection $\Gamma_1 \cap U_v$ should contain essential vertices of at most one kind is due to the fact that essential vertices have predefined distinct images in $\mathbb{P}^1$.) Our goal is to prove the converse.

5.2.2 Proposition. *Given an admissible graph Γ_0 and its admissible perturbation Γ_1, there is a deformation j_t: $B \to \mathbb{P}^1$, $t \in [0, 1]$, with the following properties:*

(1) *one has $\Gamma_0 = \Gamma(j_0)$ and $\Gamma_1 = \Gamma(j_1)$;*

(2) *the restrictions of all maps j_t to $B \smallsetminus \bigcup_v p^{-1}(U_v)$ coincide;*

(3) *the restriction of the deformation to $B \times (0, 1]$ is simple.*

Proof. Let j_0 be any ramified covering given by Theorem 5.1.7. We can assume that the restriction of j_0 to each pull-back $U'_v = p^{-1}(U_v)$ has no branch points other than the pull-backs of v itself. Then it suffices to construct a desired homotopy (fixed on the boundary) on each pull-back U'_v.

Assume that v is a $\bullet$-vertex, so that $j(v) = 0$. (In the other cases the proof is literally the same after reordering the colors and a coordinate change in $\mathbb{P}^1$.) First, assume that v is real. Let d be the full valency of v. Regard U'_v as a hemisphere in a sphere $\bar{U}'_v \cong S^2$ and extend both $\Gamma'_0 \cap U'_v$ and $\Gamma'_1 \cap U'_v$ to symmetric trichotomic graphs $\bar{\Gamma}'_0$, $\bar{\Gamma}'_1$ on $\bar{U}'_v$ by adding a real $\times$-vertex $\bar{v}$ of valency d, d $\circ$-vertices of valency 2, and appropriate edges. The graphs are admissible, and Corollary 5.1.8 gives real regular analytic maps f_0, f_1: $\bar{U}'_v = \mathbb{P}^1 \to \mathbb{P}^1$ corresponding to $\bar{\Gamma}'_0$, $\bar{\Gamma}'_1$, respectively. Clearly, $f_0(z) = z^d$ and $f_1(z)$ is a real polynomial of degree d, so that the family $f_t(z) = t^d f_1(z/t)$ is a desired homotopy. More precisely, we can assume that all critical points of f_1 other than $\bar{v}$ are mapped, say, to the disk $\{|z| < 1/2\}$ (otherwise, replace f_1 with some $\varepsilon^d f_1(\,\cdot\,/\varepsilon)$); then, $f_t^{-1}\{|z| \le 1/2\}$, $t \in I$, is a disk bundle over I, and it can be identified with $U'_v \times I$ so that the restriction of the homotopy to the boundary $\partial U'_v \times I$ is constant.

If v is not real, the same construction applies to one of the two disks constituting U'_v (with c ignored) and extends to the other disk by symmetry. □

Fix a set $\mathcal{G}$ of admissible trichotomic graphs closed under isotopies, and let $\mathcal{J}$ be the set of equivariant ramified coverings $j\colon B \to \mathbb{P}^1$ defined via $j \in \mathcal{J}$ if and only if $\Gamma(j) \in \mathcal{G}$.

5.2.3 COROLLARY. *Let $j\colon (B, c) \to (\mathbb{P}^1, {}^-)$ be an equivariant holomorphic map, $j \in \mathcal{J}$. Assume that there is a chain $\Gamma(j) = \Gamma^0, \Gamma^1, \dots, \Gamma^n$ so that $\Gamma^i \in \mathcal{G}$, $i = 0, \dots, n$, and each Γ^i, $i = 1, \dots n$, is a perturbation of, a degeneration of, or isotopic to Γ_{i-1}. Then there is a piecewise analytic equivariant deformation j_t, $t \in I$, of $j = j_0$ such that all $j_t \in \mathcal{J}$, $t \in I$, and $\Gamma(j_1) = \Gamma_n$. Moreover, each piece can be chosen as a closed real subinterval of an equivariant deformation in the sense of Kodaira-Spencer over an open complex disc.* (In general, the complex structure of B changes.)

Proof. Using Propositions 5.2.1 and 5.2.2, one can construct a topological deformation $B \times I \to \mathbb{P}^1 \times I$ as in the statement. By construction, the branch set in $\mathbb{P}^1 \times I$ can be made piecewise analytic. Moreover, by the choice made in the construction, for each (real closed) piece the equivariant ramified covering extends to an open complexification of the piece and the Grauert-Remmert theorem applies to produce a complex structure. □

5.3. Dessins. From now on, we will only consider trichotomic graphs arising from the j-invariants of almost generic elliptic surfaces (see 3.2.5) or, more generally, almost generic trigonal curves (see 3.3.4). In view of 3.3.11, this is the case if and only if

($*$) the full valency of each $\times$- (respectively, $\circ$- or $\bullet$-) vertex is 2 (respectively, 0 mod 4 or 0 mod 6).

5.3.1 PROPOSITION. *Any admissible trichotomic graph satisfying* ($*$) *above is of the form $\Gamma(j)$, where $j\colon B \to \mathbb{P}^1$ is the j-invariant of an almost generic real trigonal curve. The latter is determined uniquely up to deformation equivalence.*

Proof. The deformation uniqueness of an equivariant holomorphic map $j\colon (B, c) \to (\mathbb{P}^1, {}^-)$ such that $\Gamma = \Gamma(j)$ is given by Corollary 5.1.8. Let G_3, G_2, and I be, respectively, the sum of all $\circ$-, $\bullet$-, and $\times$-vertices considered as divisors on B. By construction, $2G_3$ is the zero divisor of j, $3G_2$ is the zero divisor of $j - 1$, and I is the pole divisor of both j and $j - 1$. In particular, $2G_3 \sim 3G_2$ and, hence, $G_2 \sim 2(G_3 - G_2)$ and $G_3 \sim 3(G_3 - G_2)$. Thus, one can take for the bundle Y generating the ruled surface (see 3.3.1) the line bundle defined by the real divisor $G_3 - G_2$.

Now, pick a real section $\tilde{g}_2 \in \Gamma(B; \mathcal{O}_B(Y^2))$ whose zero divisor is G_2 and a real section $\tilde{g}_3 \in \Gamma(B; \mathcal{O}_B(Y^3))$ whose zero divisor is G_3. For $\alpha, \beta \in \mathbb{R}$ let $g_2 = \alpha\tilde{g}_2$ and $g_3 = \beta\tilde{g}_3$. The sections $4g_2^3 j^{-1}$ and $27g_3^2(j-1)^{-1}$ of $\mathcal{O}_B(Y^6)$ are regular and have the same zero divisor I. Hence, α and β can be chosen so that j is given by (3.3.3). They are defined up to the transformation $(\alpha, \beta) \mapsto (t^2\alpha, t^3\beta)$, $t \in \mathbb{R}$; the corresponding sections g_2, g_3 define deformation equivalent trigonal curve. □

5.3.2. Any graph satisfying $(*)$ can be perturbed to a graph Γ such that

(1) the full valency of each $\times$-, $\circ$-, or $\bullet$- vertex of Γ is, respectively, 2, 4, or 6;

(2) the valency of any real monochrome vertex of Γ is 3;

(3) Γ has no inner monochrome vertices.

An admissible graph satisfying conditions (1)–(3) is called a *dessin*; such a graph corresponds to a generic trigonal curve. We always assume that the boundary of the underlying surface is nonempty. We freely extend to dessins all terminology that applies to almost generic trigonal curves. Thus, we speak about $(M-d)$-*dessins*, (*non-*)*hyperbolic* (components of) dessins, *ovals* and *zigzags* (see 5.3.6 for more details and a reinterpretation of these notions in terms of the dessins).

The ramified covering defined by a dessin Γ has generic branching behavior; its degree is of the form $6k$, $k \in \mathbb{Z}$, and one has $\#_\bullet(\Gamma) = 2k$, $\#_\circ(\Gamma) = 3k$, and $\#_\times(\Gamma) = 6k$. The number $3k$ is called the *degree* of Γ. By definition, it is positive and divisible by 3.

A dessin Γ of degree 3 on a disk is called a *cubic*. Such a dessin Γ is indeed the dessin of a nonsingular cubic curve in the projective plane blown-up at one point.

Two dessins are called *equivalent* if, after a homeomorphism of the underlying surfaces, they can be connected by a finite sequence of isotopies and the following *elementary moves*:

• *monochrome modification*, see Figure 3(a);

• *creating* (destroying) *a bridge*, see Figure 3(b), a *bridge* being a pair of monochrome vertices connected by a real monochrome edge;

• $\circ$-*in* and its inverse $\circ$-*out*, see Figure 3(c) and (d);

• $\bullet$-*in* and its inverse $\bullet$-*out*, see Figure 3(e) and (f).

(In the first two cases, a move is valid if and only if the result is again a dessin, i.e., one needs to check its admissibility.)

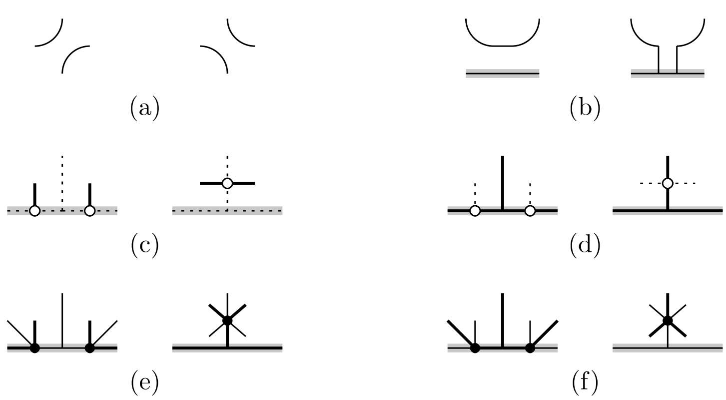


Figure 3. Elementary moves of dessins. Wide gray lines indicate real points.

Clearly, the elementary moves are exactly the results of passing through codimension 1 degenerations still satisfying ($*$). Hence, in view of Proposition 5.3.1 and Corollary 5.2.3, the following statement holds.

5.3.3 PROPOSITION. *Two generic real trigonal curves are deformation equivalent in the class of almost generic real trigonal curves if and only if their dessins are equivalent.* □

5.3.4. The definition of the j-invariant gives an easy way to recover the topology of a generic real trigonal curve $C \subset \Sigma = \mathbb{P}(1 \oplus Y)$ from its dessin Γ. Let $q\colon \Sigma \to B$ be the projection and q_C its restriction to C. The pull-back $q_C^{-1}(b) \subset q^{-1}(b)$ of each point $b \in B \smallsetminus \{\times\text{-vertices}\}$ consists of three points.

(1) If b is an inner point of a region of Γ, the three points of the pull-back $q_C^{-1}(b)$ form a triangle with all three edges distinct. As a consequence, the restriction of q_C to the interior of each region of Γ is a trivial covering.

(2) If b belongs to a dotted edge of Γ, the three points of the pull-back $q_C^{-1}(b)$ are collinear. The ratio (smallest distance)/(largest distance) is in $(0, 1/2)$; it tends to 0 ($1/2$) when b approaches a $\times$- (respectively, $\circ$-) vertex.

(3) If b belongs to a solid (bold) edge of Γ, the three points of the pull-back $q_C^{-1}(b)$ form an isosceles triangle with the angle at the vertex less than (respectively, greater than) $\pi/3$. The angle tends to 0, $\pi/3$, or π when b approaches, respectively, a $\times$-, $\bullet$-, or $\circ$-vertex.

5.3.5. In particular, statements (1)–(3) above give a very simple description of the $\mathcal{B}_3/\Delta^2$-valued monodromy, see 3.3.8, along any loop γ in $B^\sharp$. As a consequence, the following statements hold:

- if γ does not intersect the closure of Γ_{dotted}, then the monodromy along γ is determined by the corresponding permutation, which must be even (as in this case the three points in the fiber never become collinear);
- in particular, if γ does not intersect the closure of $\Gamma_{\text{dotted}} \cup \Gamma_{\text{solid}}$ (or the closure of $\Gamma_{\text{dotted}} \cup \Gamma_{\text{bold}}$), then the monodromy along γ is trivial;
- if γ belongs to the closure of Γ_{dotted}, then the monodromy along γ is $\Delta^{\epsilon \bmod 2}$, where ϵ is the number of $\circ$-vertices on γ.

5.3.6. Let $\Gamma \subset D$ be a dessin. The collection of all vertices and edges of Γ contained in a given connected component of ∂D is called a *real component* of Γ. In the drawings, (portions of) the real components of Γ are indicated by wide grey lines.

Every maximal dotted segment on a nonhyperbolic real component (respectively, every maximal real bold segment) is bounded by two $\times$- (respectively, $\bullet$-) vertices. (Here, segments are allowed to contain monochrome vertices and $\circ$-vertices.) In particular, the numbers of $\times$- and $\bullet$-vertices in each real component of Γ are even.

A real component of Γ (and the corresponding component of ∂D) is called

- *even/odd*, if it contains an even/odd number of $\circ$-vertices of Γ,
- *hyperbolic*, if all edges of this component are dotted.

In addition, define the *parity* of each maximal dotted segment of Γ and each complementary segment as the parity of the number of $\circ$-vertices contained in the segment. Equivalence of dessins preserves their even, odd, and hyperbolic components, as well as the parity of the segments. A dessin is called *hyperbolic* if all its real components are hyperbolic.

Now, let Γ be the dessin of a generic real trigonal curve $C \subset \Sigma$ (see 5.3.4 for the notation). Then, the real components Γ_i of Γ are identified with the connected components B_i of $B_{\mathbb{R}}$. The pull-back $q_C^{-1}(b)$ of a real point $b \in \partial D$ has three real points if b is a dotted point or a $\circ$-vertex adjacent to two real dotted edges; it has two real points, if b is a $\times$-vertex, and a single real point otherwise. A component Σ_i of $\Sigma_{\mathbb{R}}$ is orientable (equivalently, the restriction Y_i of $Y_{\mathbb{R}}$ is topologically trivial, see 3.3.4) if and only if the corresponding real component Γ_i is even. (Indeed, recall that Y is defined by the real divisor $G_3 - G_2$, see the proof of Proposition 5.3.1, and the restriction of G_2 to B_i is even.)

A component B_i is hyperbolic (in the sense of 3.3.4) if and only if so is Γ_i. If B_i is nonhyperbolic, its ovals and zigzags are represented by the maximal dotted segments of Γ_i, even and odd, respectively. The latter are also called *ovals* and *zigzags*. Two consecutive ovals of Γ belong to a single chain, see 3.3.4, if and only if they are separated by an even number of $\circ$-vertices.

5.4. The oval count. Let $\Gamma \subset D$ be a dessin of degree $\deg \Gamma = 3k$, and let $C \subset \Sigma$ be the corresponding trigonal curve. Its genus is $g(C) = 3k - 3\chi(D) + 1$. Introduce the following notation:

- ℓ_{even}, ℓ_{odd}: the numbers of even/odd hyperbolic real components;
- ℓ_{nh}: the number of nonhyperbolic real components;
- n_{o}, n_{z}, n_{i}: the numbers of ovals, zigzags, and inner $\times$-vertices, respectively;
- $\delta = 2 - (\ell_{\mathrm{even}} + \ell_{\mathrm{odd}} + \ell_{\mathrm{nh}}) - \chi(D)$: the "excessive" Euler characteristic.

Note that $2(n_{\mathrm{o}} + n_{\mathrm{z}} + n_{\mathrm{i}}) = 6k$ is the weighted number of $\times$-vertices. Note also that all quantities introduced are nonnegative and that $\ell_{\mathrm{nh}} > 0$ unless Γ is hyperbolic. The following statement is an immediate consequence of the discussion in 5.3.6.

5.4.1 PROPOSITION. *If Γ is an $(M - d)$-dessin, one has*

$$2\ell_{\mathrm{nh}} + \ell_{\mathrm{odd}} + n_{\mathrm{z}} + n_{\mathrm{i}} + 3\delta = d + 4. \qquad \square$$

If Γ is hyperbolic, one has $\ell_{\mathrm{nh}} = n_{\mathrm{z}} = n_{\mathrm{o}} = 0$ and $n_{\mathrm{i}} = \deg \Gamma$, and the identity in Proposition 5.4.1 takes the form

$$\ell_{\mathrm{odd}} + \deg \Gamma + 3\delta = d + 4. \tag{5.4.2}$$

As in this case one also has $\ell_{\mathrm{odd}} = \deg \Gamma \bmod 2$, the following statement holds.

5.4.3 COROLLARY. *For a hyperbolic dessin, one has $d = \delta \bmod 2$.* □

5.5. Inner ○- and ●-vertices. A dessin Γ is called *bridge free* if any bridge of Γ belongs to a monochrome real component, the latter containing exactly two vertices. A nonhyperbolic dessin Γ is called *almost connected* if each connected component of Γ contains a nonhyperbolic real component.

5.5.1 LEMMA. *Any dessin Γ is equivalent to a bridge free dessin Γ' with the same numbers of essential inner vertices. If, in addition, Γ is hyperbolic* (respectively, nonhyperbolic)*, then Γ' can be chosen connected* (respectively, almost connected).

Proof. Assume that Γ has a bridge, and denote by γ the intersection of the corresponding monochrome part of Γ and the real component containing the bridge. If γ is a whole (monochrome) component containing more than 2 vertices, pick a minimal (in the sense of $\prec$) vertex v_0, a vertex v_1 adjacent to v_0, and the other vertex $v_2 \neq v_0$ adjacent to v_1; then destroying the bridge $[v_1, v_2]$ is an admissible operation. Otherwise, γ has a bridge $[v_1, v_2]$ adjacent to an essential vertex of Γ, and destroying $[v_1, v_2]$ is also admissible.

Assume that the resulting dessin Γ is disconnected. Consider a region R whose boundary contains two circles α_1, α_2 in two different connected components Γ_1, Γ_2 of Γ. We need to show that Γ_1 and Γ_2 can be joined together provided that one of them, say Γ_2, is hyperbolic. Each of the circles α_1, α_2, has edges of all three colors. Furthermore, α_1 has a ●-vertex and, hence, an inner solid or bold edge e_1. On the other hand, all real edges of Γ_2 are dotted; hence, α_2 has an inner edge e_2 of the same color as e_1. The inner modification involving e_1 and e_2 is admissible, it does not create bridges, and it reduces the number of connected components of Γ. □

The *reduction* (a *partial reduction*) of a trichotomic graph Γ is the image $\tilde{\Gamma} \subset \tilde{D}$ of Γ in the surface $\tilde{D}$ obtained from D by contracting all (respectively, some) monochrome real components of Γ. The original graph Γ is called an *inflation* of $\tilde{\Gamma}$.

The reduction carries a natural structure of a trichotomic graph. (The image of a monochrome real component of Γ is a monochrome vertex of $\tilde{\Gamma}$ unless the resulting valency is 2; in the latter case the image is ignored and considered part of an edge. In some instances, the image of valency 2 is retained as a marked point in $\tilde{\Gamma}$.) The reduction of a bridge free dessin Γ is a dessin unless all real components of Γ are monochrome. Furthermore, if Γ is bridge free, so is its reduction. The reduction preserves the counts of inner/real essential vertices of each type. A dessin Γ is called *reduced* if it has no monochrome real components. In this (and only this) case Γ coincides with its reduction. A dessin is called *totally reduced* if it has no even real components without ×-vertices. A dessin Γ is totally reduced if and only if any dessin equivalent to Γ is reduced. A dessin is equivalent to an inflation of a totally reduced one if and only if it has an odd component or a real $\times$-vertex.

The following lemma is obvious (as $\circ$- and $\bullet$-vertices can freely be "dragged" through the marked points).

5.5.2 LEMMA. *Let Γ be a bridge free dessin, and let $\tilde{\Gamma}$ be its* (partial) *reduction. Then any dessin equivalent to $\tilde{\Gamma}$ is a partial reduction of a dessin equivalent to Γ.* $\square$

A dessin is called *peripheral* if it has no inner vertices other than $\times$-vertices.

5.5.3 PROPOSITION. *Any nonhyperbolic dessin is equivalent to a peripheral one.*

Proof. Suppose that there exists a nonhyperbolic dessin not equivalent to a dessin without inner $\circ$- and $\bullet$-vertices. Among such dessins choose a dessin Γ with the smallest number of essential inner vertices. According to Lemma 5.5.1, one can assume Γ bridge free and almost connected, and, in view of Lemma 5.5.2, it suffices to show that either Γ or its reduction $\tilde{\Gamma}$ is equivalent to a dessin with fewer inner vertices.

If all nonhyperbolic real components of Γ are monochrome, then at least one such component is adjacent to a $\bullet$-vertex, which must be inner, and a $\bullet$-out move reduces the number of essential inner vertices. Otherwise, the reduction $\tilde{\Gamma}$ is a nonhyperbolic dessin and we can replace Γ with $\tilde{\Gamma}$, i.e., assume Γ reduced. Since Γ is also bridge free, any nontrivial monochrome modification of Γ is admissible.

Define an *inner chain* (of length k) in Γ as a path $v_0, \dots, v_k$ in Γ such that all edges $[v_i, v_{i+1}]$, $0 \le i < k$, and all vertices v_i, $0 < i < k$, are inner.

First, suppose that Γ has an inner chain connecting an inner $\circ$- or $\bullet$-vertex with a nonhyperbolic real component. Let $v_0, v_1, \dots, v_k$ be a shortest inner chain with this property, and assume that either v_k is monochrome or else no inner chain of length k connects an inner $\circ$- or $\bullet$-vertex with a monochrome vertex at a nonhyperbolic real component. In particular, this assumption guarantees that the creating a bridge modifications used below in the proof are admissible. (In this proof, we are mainly interested in Γ as an abstract graph (i.e., regions do not matter), and the modifications can be performed so as to keep Condition 5.1.(8) in the definition.)

Case 0. v_k is monochrome and v_{k-1} is a $\bullet$- or $\circ$-vertex. Then the number of inner vertices is reduced by a single $\bullet$-out (respectively, $\circ$-out).

Case 1.1. v_k is a $\bullet$-vertex and v_{k-1} is a $\times$-vertex. Then $k \ge 2$ and v_{k-2} is a $\circ$-vertex. This case reduces to Case 0 by creating a bold bridge, see Figure 4.

Case 1.2. v_k is a $\bullet$-vertex and v_{k-1} is a $\circ$-vertex. Consider the region R whose boundary includes $[v_{k-1}, v_k]$ and the inner solid edge incident to v_k, see Figure 5. The vertex u following v_k, v_{k-1} in the boundary of R is a $\times$-vertex. If necessary, reduce R to a triangle by a monochrome modification. Then creating a bold bridge reduces this case to Case 0.

Case 2.1. v_k is a $\circ$-vertex and v_{k-1} is a $\times$-vertex. Then $k \ge 2$ and v_{k-2} is a $\bullet$-vertex, and creating a bold bridge reduces this case to Case 0, see Figure 6.

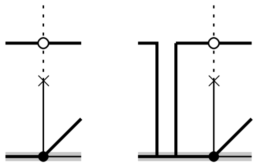


Figure 4.

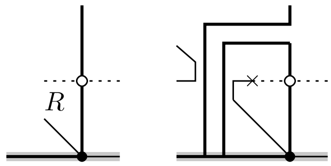


Figure 5.

Case 2.2. v_k is a $\circ$-vertex and v_{k-1} is a $\bullet$-vertex. If among the real neighbors of v_k (i.e., real vertices connected to v_k by a real edge) there is a $\times$-vertex, creating a solid bridge reduces this case to Case 0, see Figure 7. Otherwise, the real neighbors of v_k are monochrome. Let a be one of them, and let w be the $\circ$-vertex following a in the real component. Since the real component is nonhyperbolic, w is distinct from v_k. Consider the region R whose boundary includes $[v_k, v_{k-1}]$ and $[v_k, a]$, see Figure 8. The vertex u following v_k, v_{k-1} in the boundary of R is a $\times$-vertex. If necessary, reduce R to a triangle by a monochrome modification and, if v_{k-1} and w are not adjacent, perform a monochrome modification to create a bold edge $[v_{k-1}, w]$, see Figure 8. Now, replace the original chain with $v_0, \ldots,$ v_{k-1}, $v'_k = w$. Since the real component in question is nonhyperbolic, iterating this procedure (in the same direction) will produce a chain $\ldots$, v_{k-1}, v''_k with v''_k having a $\times$-vertex as a real neighbor. This reduces the situation to that considered at the beginning of this paragraph (Figure 7).

Case 3. v_k is monochrome and v_{k-1} is a $\times$-vertex. Then $k \geq 2$ and v_{k-2} is a $\circ$- or $\bullet$-vertex. By a monochrome modification one can create a bold edge connecting v_{k-2} with one of the real neighbors of v_k and thus reduce this case to Case 1.2 (see Figure 9) or 2.2 (see Figure 10).

Now, suppose that Γ has no inner chain connecting an inner $\circ$- or $\bullet$-vertex to a nonhyperbolic real component. Note that any inner chain connecting two hyperbolic real components has a $\bullet$-vertex. Since Γ is almost connected, one can find two inner chains $C = (v_0, \ldots, v_k)$ and $C' = (v'_0, v'_1, \ldots)$ so that v_k belongs to a nonhyperbolic real component, v_0 and v'_0 are connected by a real edge in a hyperbolic real component, and C' contains an inner $\circ$- or $\bullet$-vertex. Observe that $k = 1$ or 2, in the latter case v_1 being a $\times$-vertex. Denote by R the region incident to $[v_0, v'_0]$.

Case 4. v'_0 is a $\circ$-vertex. Then v_0 is monochrome and v'_1 is an inner $\bullet$-vertex. If $k = 1$, then v_1 is monochrome, the vertex following v_0, v_1 in the boundary

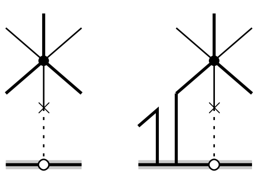


Figure 6.

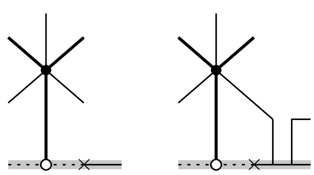


Figure 7.

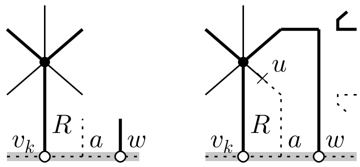


Figure 8.

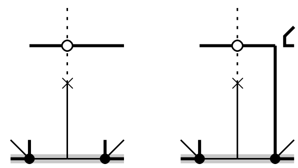


Figure 9.

of R is a real $\times$-vertex, and creating a solid bridge reduces this case to Case 0, see Figure 11. If $k = 2$, the reduction to Case 0 is obtained by creating a solid bridge as in Figure 12 (if v_2 is monochrome) or by creating a bold bridge as in Figures 14 and 13 (if v_2 is a $\bullet$-vertex and the bold edge following v_2 in the boundary of R is, respectively, inner or real; in the former case, a solid inner modification is performed first).

Case 5. v_0' is monochrome. Then v_1' is a $\times$-vertex, v_2' is a $\bullet$-vertex, $k = 1$, and v_0 is a $\circ$-vertex. This case is reduced to Case 0 by creating a bold bridge as in Figure 15 (if v_1 is monochrome) or by creating a solid bridge as in Figures 16 and 17 (if v_1 is a $\bullet$-vertex and the solid edge following v_1 in the boundary of R is, respectively, real or inner; in the latter case, a bold inner modification is required). $\square$

The next statement is an analogue of Proposition 5.5.3 for hyperbolic dessins.

5.5.4 Proposition. *Any hyperbolic dessin is equivalent to a dessin whose all $\circ$-vertices are real.*

Proof. As in Proposition 5.5.3, one can assume the dessin Γ in question bridge free, connected, and reduced. Consider a shortest inner chain $v_0, \ldots, v_k$ connecting an inner $\circ$-vertex v_0 with a real vertex v_k. It is easy to see that $k \leq 3$ and, since Γ is bridge free, $k > 1$.

If $k = 2$, then v_1 is a $\bullet$-vertex and v_2 is a $\circ$-vertex, see Figure 18. Consider the region R as in the figure and, if necessary, reduce it to a triangle by a monochrome modification. Now, the number of inner $\circ$-vertices is reduced by creating a dotted bridge followed by a $\circ$-out, see Figure 18.

If $k = 3$, then either v_1 is a $\times$-vertex, v_2 is a $\bullet$-vertex, and v_3 is a $\circ$-vertex (see Figure 19), or v_1 is a $\bullet$-vertex, v_2 is a $\times$-vertex, and v_3 is monochrome (see Figure 20). In the former case, all three $\circ$-vertices adjacent to v_2 are real

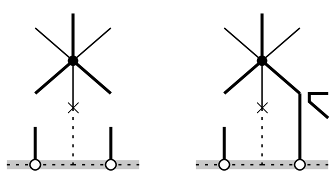


Figure 10.

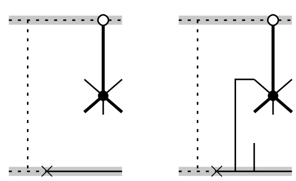


Figure 11.

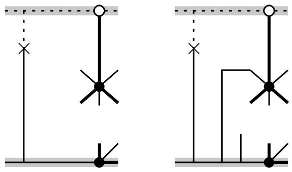


Figure 12.

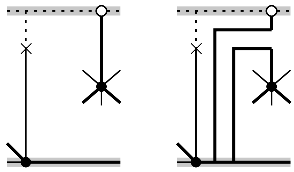


Figure 13.

(as otherwise the chain $v_0, \ldots, v_k$ would not be shortest), and the number of inner $\circ$-vertices is reduced by creating a dotted bridge followed by a $\circ$-out, see Figure 19. In the latter case, all three $\circ$-vertices adjacent to v_1 are inner, and at least one of them (not necessarily v_0) can be pushed out by creating a dotted bridge followed by a $\circ$-out, see Figure 20. $\square$

5.6. Indecomposable dessins. In this section, we allow dessins on disconnected surfaces (which are merely unions of dessins on the components of the surface).

Consider a dessin $\Gamma \subset D$. Let $I_1, I_2 \subset \partial D$ be a pair of segments whose endpoints are not vertices of Γ, and let $\varphi\colon I_1 \to I_2$ be an isomorphism, i.e., a diffeomorphism of the segments establishing a graph isomorphism $\Gamma \cap I_1 \to \Gamma \cap I_2$ and preserving the kinds of the vertices and edges. (Note that, if I_1 contains at least one essential vertex of Γ, then φ necessarily preserves the orientations of the edges given by the trichotomic graph structure.) Consider the quotient $D_\varphi = D/\{x \sim \varphi(x)\}$ and the image $\Gamma'_\varphi \subset D_\varphi$ of Γ, and denote by Γ_φ the graph obtained from Γ'_φ by erasing the image of I_1, if φ is orientation reversing, or converting the images of the endpoints of I_1 to monochrome vertices otherwise.

In what follows we always assume that either I_1 is part of an edge of Γ or I_1 contains a single $\circ$- or $\times$-vertex. In the latter case, φ is unique up to isotopy; in the former case, φ is determined by whether it is orientation preserving or orientation reversing. If Γ_φ is a dessin, it is called the result of *gluing* Γ along φ. (Sometimes we speak about gluing several dessins, meaning gluing their disjoint union.) The image of I_1 is called a *cut* in Γ_φ, and Γ is called the result of a cut. The cut is called *genuine* (*artificial*) if φ is orientation preserving (respectively, reversing); it is called a solid, dotted, bold, or $\times$-cut according to the structure of $\Gamma \cap I_1$. (The terms dotted and bold still apply to cuts containing a $\circ$-vertex.)

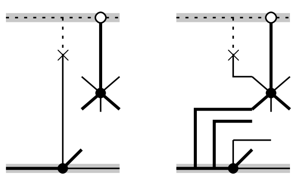


Figure 14.

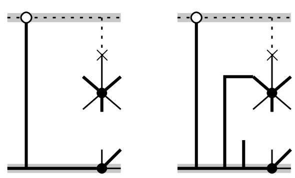


Figure 15.

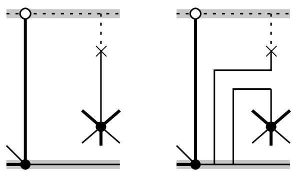


Figure 16.

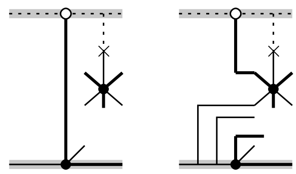


Figure 17.

A dessin that is not equivalent to the result of gluing another dessin is called *indecomposable*. A *generalized cubic* is a dessin whose reduction is a cubic.

5.6.1 THEOREM. *Any indecomposable dessin is a disjoint union of generalized cubics.*

5.6.2 COROLLARY. *Any dessin can be obtained from a disjoint union of generalized cubics by a sequence of gluing operations and equivalences.* □

Remark. At present, we do not know whether a given graph is equivalent to the result of gluing of a union of cubics. As shown below, this is true for *M*- and ($M - 1$)-dessins.

In view of Propositions 5.5.3 and 5.5.4, Theorem 5.6.1 is an immediate consequence of Propositions 5.6.3 (the hyperbolic case) and 5.6.4 (the nonhyperbolic case).

5.6.3 PROPOSITION. *Let* Γ *be a connected reduced hyperbolic dessin whose all* $\circ$*-vertices are real. Then* Γ *either is a cubic, or has a cut; in the former case,* Γ *is isotopic to the dessin shown in Figure 22.*

Proof. Consider a $\bullet$-vertex v of Γ. Under the hypothesis, v has a neighborhood shown in Figure 21. If this neighborhood does not close up to a cubic (*i.e.*, at least one of the regions adjacent to v is not a triangle), then Γ has an artificial dotted cut (located in the above region). □

Remark. One can show that, on the disc, any two hyperbolic dessins of the same degree are equivalent. If all $\circ$-vertices are real, such a dessin Γ is a perturbation of a star-like trichotomic graph as in Figure 22, with $2\deg\Gamma$ alternating rays radiating from a single multiple $\bullet$-vertex. (Note that the latter graph does satisfy 5.3($*$), and thus represents the *j*-invariant of an almost generic curve, see Proposition 5.3.1.)

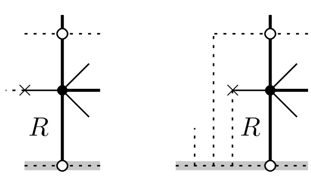


Figure 18.

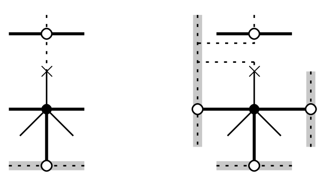


Figure 19.

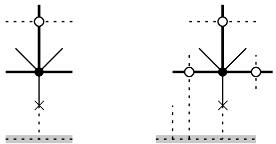


Figure 20.

5.6.4 Proposition. *Let Γ be a reduced peripheral dessin on a connected surface. Then either Γ is a cubic, or Γ is equivalent to a peripheral dessin with a cut.*

Proposition 5.6.4 is a mere combination of Lemmas 5.6.6 and 5.6.7 proved at the end of this section.

Given a region R, a component of the boundary ∂R is called a *$3m$-gonal component* if it contains $3m$ essential vertices (equivalently, m vertices of any given kind). If ∂R consists of a single $3m$-gonal component, then R itself is called a *$3m$-gon*.

Recall that the real $\circ$-vertices of a dessin can be subdivided into two types, depending on the type of the real edges incident to the vertex. Similarly, the real $\bullet$-vertices in the boundary of a given region R can be subdivided into three types, depending on which of the three angles at the vertex belongs to R.

5.6.5 Lemma. *Let R be a region in a reduced peripheral indecomposable dessin. Then the following holds:*

(1) *the boundary ∂R cannot contain two distinct real edges of the same kind;*

(2) *the boundary ∂R cannot contain two distinct $\circ$-vertices of the same type;*

(3) *the boundary ∂R cannot contain two distinct $\bullet$-vertices of the same type;*

(4) *the boundary ∂R consists of either one or two triangles or a hexagon;*

(5) *unless R is a triangle, the boundary ∂R cannot contain an inner $\times$-vertex adjacent to a solid monochrome vertex;*

(6) *if ∂R is disconnected, it cannot contain a real $\times$-vertex.*

Proof. If ∂R contains two real edges of the same kind, they either are connected by an inner edge of the same kind or can be connected by an artificial cut; in both cases the graph is decomposable. This proves (1). Statement (2) follows

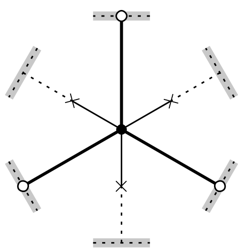


Figure 21.

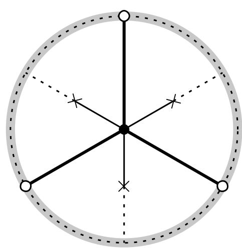


Figure 22.

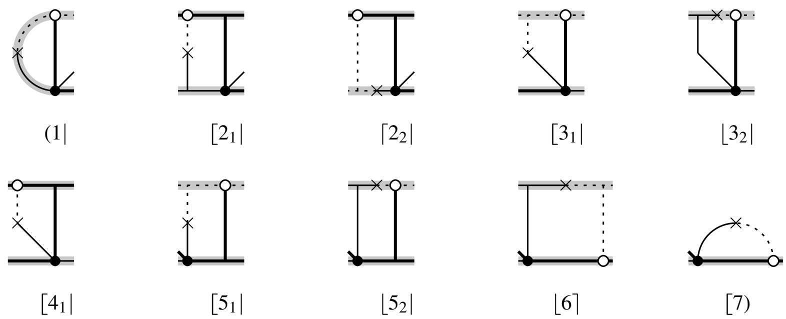


Figure 23. Triangular regions of indecomposable dessins.

directly from (1), and (3) follows from (1) unless R has no real edges at the two vertices in question. In the latter case, a bold inner modification results in a region with two distinct solid real edges, which contradicts (1). (Alternatively, a solid inner modification results in a region with two distinct bold real edges.)

In view of (2), ∂R contains at most two $\circ$-vertices. This implies (4).

Let u be a $\times$-vertex as in (5). In view of (2), since ∂R is not a triangle, it contains a $\circ$-vertex incident to dotted real edges. Then, creating a dotted bridge produces a cut (containing u).

Let u be a real $\times$-vertex in ∂R and let v be a $\times$-vertex in another component of ∂R. Due to (1), v is an inner vertex, and one can create a solid bridge (close to u), converting R to a hexagon and v, to a $\times$-vertex as in (5). $\square$

In Lemma 5.6.6 below we list all regions appearing in an indecomposable dessin (see Figures 23 and 24). Various brackets in the notation indicate the "ends" of a region, i.e., the components of the inner parts of its boundary. (Clearly, it is these components that govern the adjacencies of the regions.) The symbols $|$, $\lfloor$, and $\lceil$ (and the corresponding right delimiters) stand, respectively, for a bold, solid, and dotted edge, and $[$ stands for a pair of edges separated by an inner $\times$-vertex. The brace $\{$ indicates several "ends" that are not of particular interest, and (indicates no "end" at all.

5.6.6 LEMMA. *Any region R in a reduced peripheral indecomposable dessin is either one of the triangles in Figure 23 or one of the hexagons in Figure 24.*

Proof. Lemma 5.6.5(1) restricts all possible triangle components of the boundary of R to those listed in Figures 23 and 25, and (1)–(3) and (5) restrict the hexagons to those listed in Figure 24. Furthermore, a hexagon bounds a region,

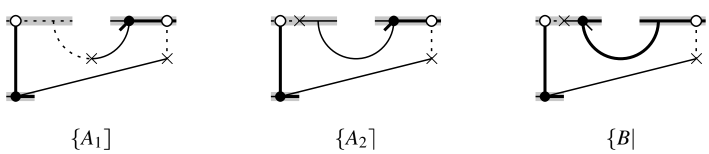


Figure 24. Hexagonal regions of indecomposable dessins.

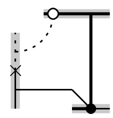


Figure 25. The exceptional triangle.

see 5.6.5(4), and if the latter is not a disk, it can be modified to a region with disconnected boundary, see below.

Assume that R is the triangle in Figure 25. Its bold edge can only be adjacent to a triangle of type 2_1 or 2_2 or a hexagon of type B. In the former case, a $\circ$-in modification followed by a $\circ$-out along any dotted edge produces a dotted cut. In the latter case, a bold inner modification within the hexagon results in a region with two solid (as well as two dotted) real edges.

Finally, assume that ∂R consists of two triangles. Lemma 5.6.5(6) reduces the list of triangles to 2_1, 3_1, 4_1, 5_1, and 7, and 5.6.5(5) eliminates 2_1. Thus, in view of 5.6.5(1), the boundary ∂R must be formed by one of the pairs 3_1, 4_1 or 3_1, 7. The former is eliminated by 5.6.5(3), and in the latter case, a solid (or dotted) inner modification results in a region with two bold real edges. $\square$

5.6.7 Lemma. *Any reduced dessin* (on a connected surface) *whose regions are those listed in Lemma 5.6.6 is a cubic. Conversely, all regions of a peripheral cubic are among those listed in Lemma 5.6.6, and they are attached to one another according to one of the following adjacency schemes (see Figure 26):*

$$
\begin{array}{rccccccccccc}
\mathrm{II}_1: & (1| & — & |3_1] & — & [5_1| & — & |5_1] & — & [3_1| & — & |1) \\
\mathrm{I}_1: & (1| & — & |3_1] & — & [5_1| & — & |5_2\rfloor & — & \lfloor 3_2| & — & |1) \\
\mathrm{I}_1: & (1| & — & |3_2\rfloor & — & \lfloor 5_2| & — & |5_1] & — & [3_1| & — & |1) \\
\mathrm{II}_3: & (1| & — & |3_2\rfloor & — & \lfloor 5_2| & — & |5_2\rfloor & — & \lfloor 3_2| & — & |1) \\
\mathrm{I}_2: & (1| & — & |3_2\rfloor & — & \lfloor 6\rceil & — & \lceil 6\rfloor & — & \lfloor 3_2| & — & |1) \\
\mathrm{I}_1: & (1| & — & |3_2\rfloor & — & \lfloor 6\rceil & — & \lceil 2_2| & — & |4_1] & — & [7) \\
\mathrm{II}_0: & (7] & — & [4_1| & — & |2_1] & — & [2_1| & — & |4_1] & — & [7) \\
\mathrm{I}_0: & (7] & — & [4_1| & — & |2_2\rceil & — & \lceil 2_2| & — & |4_1] & — & [7)
\end{array}
$$

$$
\begin{array}{rccccc}
\mathrm{I}_0: & \{A_1] & — & [3_1| & — & |1) \\
\mathrm{II}_2: & \{A_2\rfloor & — & \lfloor 3_2| & — & |1) \\
\mathrm{II}_1: & \{B| & — & |4_1] & — & [7)
\end{array}
$$

(in the last three cases each hexagon being also adjacent to a triangle of type 1 *and a triangle of type* 7*).*

Remark. Some pairs of dessins listed in Lemma 5.6.7 and Figure 26 are equivalent. It is easy to see that, in fact, there are seven equivalence classes of cubics. They differ by the type (I or II; equivalently, cubics with an oval are of type I, and those without ovals are of type II) and the number of zigzags (shown as a subscript in the notation). The equivalence class represented by each dessin is also listed in Lemma 5.6.7 and Figure 26.

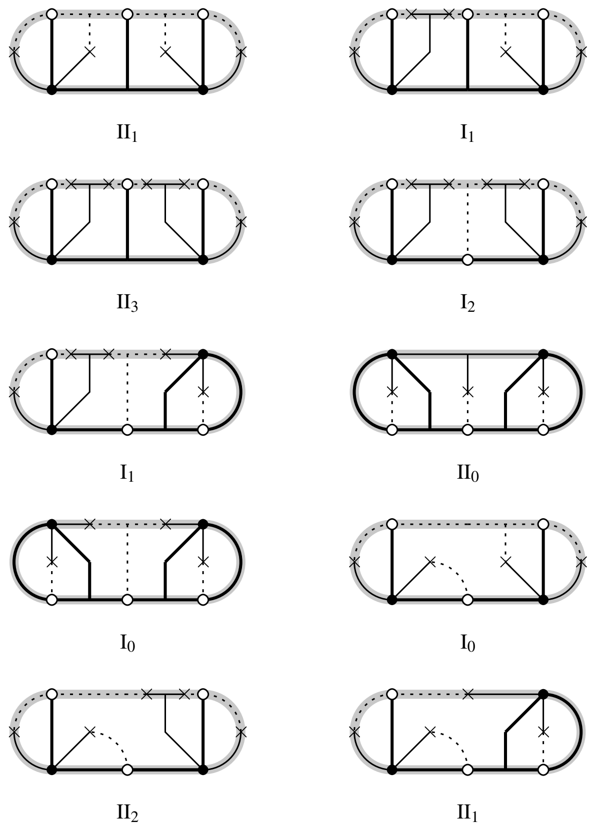


Figure 26. Peripheral cubic dessins.

Proof. It suffices to consider a dessin whose all regions are among those listed in Lemma 5.6.6. (Any reduced nonhyperbolic cubic has this property since it is indecomposable.) Comparing the "ends" of the regions, one arrives at the following list of adjacencies:

$$\begin{array}{rcccl}
 & & (1\rvert & — & \lvert 3_1], \lvert 3_2\rfloor, \{A_1], \{A_2\rfloor, \{B\rvert \\
\lvert 2_1] & — & [2_1\rvert & — & \lvert 4_1] \\
\lvert 2_2\rceil, \lfloor 6\rceil & — & \lceil 2_2\rvert & — & \lvert 4_1] \\
\lvert 5_1], \{A_1] & — & [3_1\rvert & — & \lvert 1) \\
\lvert 5_2\rfloor, \lceil 6\rfloor, \{A_2\rfloor & — & \lfloor 3_2\rvert & — & \lvert 1) \\
(7] & — & [4_1\rvert & — & \lvert 2_1], \lvert 2_2\rceil, \lvert B\} \\
\lvert 3_1] & — & [5_1\rvert & — & \lvert 5_1], \lvert 5_2\rfloor \\
\lvert 3_2\rfloor & — & \lfloor 5_2\rvert & — & \lvert 5_1], \lvert 5_2\rfloor \\
\lvert 3_2\rfloor & — & \lfloor 6\rceil & — & \lceil 2_2\rvert, \lceil 6\rfloor \\
 & & (7] & — & [4_1\rvert, \{A_1], \{A_2\rfloor, \{B\rvert.
\end{array}$$

It remains to list all chains of regions joined according to these rules, terminating a chain whenever there are no free "ends" left.

Assume that all regions of Γ are triangles. If Γ has a triangle of type 1, starting from it one obtains one of the first six schemes in the statement. Otherwise, Γ has no triangle of types 3_1, 3_2 and, hence, no triangle of types 5_1, 5_2, or 6. Assuming that Γ has a triangle of type 7, one arrives at the last two schemes with triangles only. Otherwise, Γ has no triangle of type 4_1 and, hence, no triangle of type 2_1 or 2_2, i.e., such a dessin does not exist. Finally, any hexagon that Γ may have extends uniquely to one of the last three schemes in the statement. It is straightforward to observe that all eleven schemes do represent cubics. $\square$

5.7. Scraps. Given a dessin and one or several of its inner edges, each connecting a real $\circ$-vertex and a real monochrome vertex, one can cut the dessin along these edges; the connected components of the result (which, in general, is not a dessin anymore) are called *scraps*. The edges used in the cut are called *breaks*; they can be dotted or bold. (In the sequel we need dotted breaks only.) Note that a scrap with breaks is not a dessin; it can be regarded as a "dessin with boundary." Two scraps can be glued along a break of the same kind. The result is a dessin if and only if it is admissible and has no breaks.

We extend to scraps the weighted numbers $\#_\circ$, $\#_\bullet$, and $\#_\times$. Given a scrap σ on a surface D, denote by $\beta(\sigma)$ the number of breaks in the boundary of σ. Let further $\kappa(\sigma) = \chi(D) - \frac{1}{2}\beta(\sigma)$. The latter quantity is additive; one has $\kappa(\sigma) > 0$ if and only if D is a disk and $\beta(\sigma) \leq 1$, and $\kappa(\sigma) = 0$ if and only if D is a disk and $\beta(\sigma) = 2$, or D is an annulus or a Möbius band and $\beta(\sigma) = 0$. Another additive quantity associated to a scrap σ is the degree $\deg(\sigma) = \#_\circ(\sigma) - \frac{1}{2}\beta(\sigma)$. The degree of any scrap is positive.

5.7.1 LEMMA. *For a scrap σ, one has* $\#_\times(\sigma) = 2\deg(\sigma)$ *and* $\#_\bullet(\sigma) = \frac{2}{3}\deg(\sigma)$. *Furthermore,* $\deg(\sigma) + \frac{3}{2}\beta(\sigma) = 0 \bmod 3\mathbb{Z}$.

Proof. It suffices to complete σ to a true dessin by patching each break with a half of a cubic (say, $\lfloor 5_1\rceil$—$\lceil 3_1|$—$|1)$ or $\lceil 6\rfloor$—$\lfloor 3_2|$—$|1)$ in the notation of section 5.6) and to use the known identities and congruences for dessins. $\square$

5.7.2 COROLLARY. *A scrap σ with* $\beta(\sigma) = 1$ (respectively, 2) *has* $\deg(\sigma) \geq \frac{3}{2}$ (respectively, $\deg(\sigma) \geq 3$). $\square$

5.7.3. The importance of scraps is in the following construction. Let Γ be a dessin (or a scrap). Each oval and each odd real hyperbolic component of Γ has at least one dotted monochrome vertex u. Let e be the inner edge incident to u (and extended through any inner $\circ$-vertex), and let v be the other end of e. Then either v is an inner $\times$-vertex, or v is a monochrome vertex and, hence, e is a dotted cut, or else v is a real $\circ$-vertex. In the last case, e has no inner vertices, and, thus, breaks Γ into smaller scrap(s).

As an immediate consequence, since a monochrome vertex u in an odd real hyperbolic bridge free component cannot be adjacent to a $\circ$-vertex not in the component, we obtain the following statement.

5.7.4 LEMMA. *If a dessin* Γ *has no genuine dotted cuts, then* $\ell_{\mathrm{odd}} \le n_{\mathrm{i}}$. □

5.7.5 LEMMA. *A scrap* σ *with* $\kappa(\sigma) > 0$ *contains a zigzag or an inner* $\times$*-vertex.*

Proof. If σ has no inner $\times$-vertices, one can use 5.7.3 to subdivide it into smaller scraps so that none of them has ovals. At least one of the pieces still has $\kappa > 0$. Such a piece σ' can only be a scrap on a disk with $\beta(\sigma') \le 1$. Due to Lemmas 5.7.2 and 5.7.1 it has at least three $\times$-vertices and, hence, at least one zigzag. □

5.7.6 THEOREM. *If an* $(M-d)$*-dessin* Γ *has no genuine dotted cut, then*

$$2\deg\Gamma \le 3(n_{\mathrm{z}} + n_{\mathrm{i}}) + 3d - 3\delta.$$

Proof. Let $\deg\Gamma = 3k$. Using the construction of 5.7.3, one can break Γ into scraps, the total number of breaks being $2b$, where

$$b \ge b_0 = \ell_{\mathrm{odd}} + n_{\mathrm{o}} - n_{\mathrm{i}} = \ell_{\mathrm{odd}} + 3k - n_{\mathrm{z}} - 2n_{\mathrm{i}} \tag{5.7.7}$$

(we count each break twice, once in each of the two scraps incident to it). Let m_+ be the nubmer of scraps σ with $\kappa(\sigma) > 0$. Using Lemma 5.7.1 one can split $m_+ = m'_+ + m''_+$, where m'_+ is the number of scraps with $\deg(\sigma) = \frac{3}{2}$ and m''_+ is the number of scraps with $\deg(\sigma) \ge \frac{9}{2}$. According to Lemma 5.7.5, at least $m'_+ - n_{\mathrm{z}}$ inner $\times$-vertices are separated by breaks from the ovals, and the inequality (5.7.7) can be sharpened to $b \ge b_0 + m'_+ - n_{\mathrm{z}}$.

Let b_- be the total number of breaks in the scraps with $\beta \ge 3$. Then, according to Corollary 5.7.2 and the definition of m''_+, the number $3k$ of $\circ$-vertices of Γ is at least $b+(2b-b_-)+3m''_+ \ge 3\ell_{\mathrm{odd}}+9k-6(n_{\mathrm{z}}+n_{\mathrm{i}})-b_-+3m_+$. Hence, one must have $6k \le 6(n_{\mathrm{z}}+n_{\mathrm{i}})-3\ell_{\mathrm{odd}}+b_- -3m_+$. On the other hand, since $\kappa/\beta \le -\frac{1}{6}$ for a scrap with $\beta \ge 3$, the additivity of κ yields $\frac{1}{6}b_- \le \frac{1}{2}m_+ - \chi(D)$. Hence, $6k \le 6(n_z+n_i)-3\ell_{\mathrm{odd}}-6\chi(D)$, and it remains to substitute $\chi(D) = 2-(\ell_{\mathrm{odd}}+\ell_{\mathrm{nh}})-\delta$ and use Proposition 5.4.1. □

6. Applications: *M*- and (*M* − 1)-cases. Recall that we only consider generic curves and surfaces and classify them up to deformation in the class of almost generic ones. The base of the fibration is never assumed fixed; it is also subject to a deformation. Unless stated otherwise, trigonal curves never intersect the exceptional section.

6.1. Junctions. Define a (*self-*)*junction* as a genuine gluing of a dessin along isomorphic parts of two zigzags (respectively, one zigzag) so that the resulting

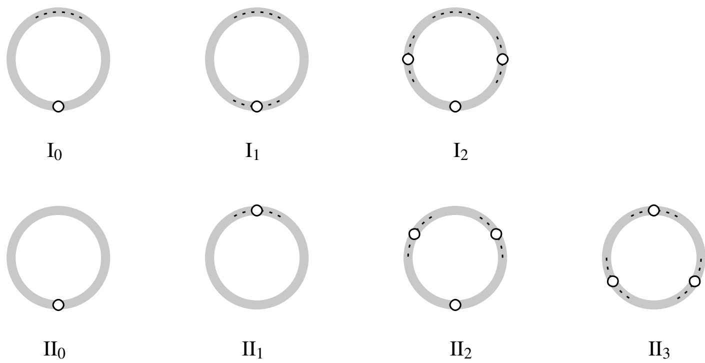


Figure 27. Ribbon boxes; odd segments are marked with ∘-points.

cut connects two ovals (respectively, an oval and an odd hyperbolic component) of the dessin obtained, see 5.6 for the terminology concerning cuts. Note that a (self-)junction consumes the zigzags involved. In particular, any two junctions commute.

Below we state several structure theorems that deal with junctions of cubics. From the point of view of the junction operation, a cubic can be regarded as a "black box" with a certain extra decoration of its boundary. More precisely, we define a *ribbon box* as a disc with a few disjoint segments (called *dotted*) marked in the boundary, and a parity assigned to each dotted and each complementary segment; a box is required to be one of those listed in Figure 27.

Each nonhyperbolic cubic dessin Γ gives rise to a ribbon box: the disk is the underlying surface of Γ, the dotted segments of the box are the maximal real dotted segments of Γ, and the parity is given by the number of ∘-vertices, as in the cubic, cf. 5.3.6. Conversely, in view of the classification given by Lemma 5.6.7, each box is obtained in this way from a cubic dessin which is unique up to equivalence leaving real ×-vertices fixed. In view of this correspondence, we will refer to even (odd) dotted segments of a ribbon box as oval (respectively, zigzag) segments.

A *ribbon curve structure* is a collection of boxes in which some of the boxes are glued via identifying certain pairs of zigzag segments, so that the result is a connected surface, and some of the remaining zigzag segments are selected for future self-junction. The selected segments are called *vanishing*. An *isomorphism* of two ribbon curve structures is a homeomorphism of the underlying surfaces preserving the decorations, *i.e.*, taking boxes to boxes and dotted segments to dotted segments, preserving all parities, and taking vanishing segments to vanishing segments.

Each of the two zigzag segments of a box of type I_2 or II_2 can be given a preferred orientation, say, towards its odd complementary neighbor. Thus, each adjacency of two such boxes has a sign: it is said to be *positive* or *negative* depending on whether the orientations of two segments involved do or do not coincide.

The *junction graph* of a ribbon curve structure is the graph obtained by replacing each box with a vertex and connecting each pair of glued boxes by an edge (one for each pair of segments identified). The junction graphs of isomorphic ribbon curve structures are isomorphic. Clearly, the valency of a vertex of a junction graph does not exceed the number of zigzag segments of the box represented by the vertex. In particular, the valency is at most 3, and any vertex of valency 3 represents a box of type II_3.

A ribbon curve structure defines an equivalence class of totally reduced dessins: one replaces each box with a corresponding cubic, performs a junction on each pair of identified zigzag segments, and performs a self-junction on each vanishing segment. A *ribbon curve* is a trigonal curve whose dessin is equivalent to one obtained in this way. Note that, since cubics are indecomposable, each particular dessin admits at most one ribbon curve structure.

An *enhanced ribbon curve structure* is a ribbon curve structure equipped with a collection of nonnegative integers, one for each box other than II_3, one for each adjacency, and one for each vanishing segment. The notion of isomorphism extends naturally: one requires that the isomorphism should preserve the enhancement. An enhanced ribbon curve structure defines an equivalence class of dessins: one takes the totally reduced dessin constructed in the previous paragraph and inflates it by placing the indicated number of dotted monochrome components to each (self-)junction and to an inner dotted edge within each box. Considering the types of boxes one by one, see Lemma 5.6.7, one can easily show that the equivalence class is indeed well defined: for each cubic dessin, any two distinct inner dotted edges in it can be connected by a sequence of elementary moves; hence, the dotted monochrome components introduced inside the box can be placed to any preselected inner dotted edge.

6.2. Classification of trigonal M-curves.

6.2.1 THEOREM. *The collection of all dotted cuts of any totally reduced nonhyperbolic M-dessin Γ represents Γ as an iterated (self-)junction of a union of M-cubics. Furthermore, any elementary move of Γ is either a simple modification of the junction or an elementary move in one of the cubics* (not involving the cuts).

Proof. We will show that a totally reduced nonhyperbolic M-dessin $\Gamma \subset D$ that is not a cubic is a (self-)junction of another M-dessin. One has $d = 0$; hence, $\delta = 0$ and $n_z + n_i \leq 2$ if D is a disk, and $n_z + n_i \leq 1$ otherwise, see 5.4. Thus, according to Theorem 5.7.6, the dessin Γ has a dotted cut. From the oval count given by Proposition 5.4.1 it follows that any such cut is a (self-)junction, the result being an M-dessin. (Roughly, to keep the maximal number of components, each gluing must create at least two ovals, and each self-gluing must create at least three components. The possibility to form an even hyperbolic component is ruled out by the assumption that the dessin is totally reduced.) In particular, the result of the cut has no hyperbolic components and, hence, is still totally reduced.

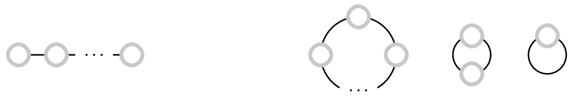


Figure 28. Junction graphs of M-dessins.

The second statement follows from the first one and the fact that a cubic is indecomposable. Indeed, the only elementary move not as in the theorem is an inner modification joining two cuts (within one cubic) and producing an alternative pair of cuts. However, from the point of view of the cubic, that would imply the existence of an artificial dotted cut, which would contradict to the fact that the cubic is indecomposable. □

6.2.2 COROLLARY. *Any totally reduced nonhyperbolic M-dessin admits a ribbon curve structure, which has the following properties:*

- *each box represents an M-cubic;*
- *the underlying surface is orientable.*

Conversely, any ribbon curve structure with the properties above defines a totally reduced M-dessin. Furthermore, two such dessins are equivalent if and only if their ribbon curve structures are isomorphic.

Remark. As each box representing an M-cubic has valency at most 2, the junction graph of an M-curve is either a linear tree or a single cycle, see Figure 28. In the former case, the underlying surface is a disk, in the latter case it is an annulus.

Remark. The statement of Corollary 6.2.2 for the case of rational base was first obtained by S. Orevkov [Or2].

Proof of Corollary 6.2.2. The only statement that needs proof is the fact that, if the underlying graph is a single cycle, the underlying surface cannot be a Möbius band. This possibility is eliminated by Proposition 5.4.1. □

6.2.3 THEOREM. *The deformation classes of almost generic nonhyperbolic trigonal M-curves are in a canonical one-to-one correspondence with the isomorphism classes of enhanced ribbon curve structures as in Corollary 6.2.2.*

Proof. Due to Proposition 5.3.3, it suffices to enumerate the equivalence classes of nonhyperbolic M-dessins. Any such dessin Γ has ovals, and hence real $\times$-vertices. Thus, Γ is equivalent to an inflation of a totally reduced dessin, and, in view of Lemma 5.5.2, the statement of the theorem follows from Corollary 6.2.2. □

6.2.4 THEOREM. *For each integer $g \geq 0$, there is a unique deformation class of almost generic hyperbolic trigonal M-curves over a base of genus g.*

Proof. Proposition 5.3.3 reduces the problem to dessins. Consider a hyperbolic M-dessin. The oval count (5.4.2) implies that $\delta = 0$, $\ell_{\mathrm{odd}} = 1$, and $\deg \Gamma = 3$. Hence, Γ is equivalent to an inflation of a totally reduced dessin, which is a hyperbolic cubic; the latter is equivalent to the dessin shown in Figure 22, see Lemma 5.5.4. As above, the equivalence class of the inflation is determined by the number g of the components inserted. □

6.3. Classification of elliptic *M*-surfaces. Define a *ribbon surface structure* as a ribbon curve structure satisfying the following additional requirements:

- there are no vanishing segments;
- the combined parity of the segments within each boundary component of the underlying surface is even;

and enriched with the following decorations:

- each junction (i.e., common segment of two boxes glued together) is subdivided into segments, and each segment and each inner vertex of the subdivision are given a sign;
- each box other than II_3 is given a sequence of signs of even length.

The decorations must be subject to the following condition: the product of the signs given to all segments in the junctions adjacent to one box is -1.

From the definition it follows that the total number of boxes in a ribbon surface structure must be even. In particular, there are no boxes of types I_0 or II_0.

An *isomorphism* of ribbon surface structures is an isomorphism of the corresponding ribbon curve structures preserving the additional decorations.

6.3.1 THEOREM. *Each almost generic elliptic M-surface defines an isomorphism class of ribbon surface structures with the following properties:*

- *each box represents an M-cubic;*
- *the underlying surface is orientable.*

(cf. Corollary 6.2.2). Conversely, each ribbon surface structure with the properties above defines one or two (depending on whether the underlying surface is a disk or an annulus, respectively) *deformation classes of pairs of opposite elliptic M-surfaces.*

Proof. Essentially, the statement follows from Theorem 6.2.3; we will just explain the relation between elliptic surfaces, trigonal curves, ribbon surface structures, and enhanced ribbon curve structures. In view of Corollary 4.3.7, up to deformation each M-surface is Jacobian; hence, it is described by its Weierstraß model, see 3.3.5, and the branch curve is M-, see Lemma 3.3.7. Thus, an M-surface determines and is determined by the following data: a trigonal M-curve C over the same base B (hence, an enhanced ribbon curve structure), a lift of the monodromy $\pi_1(B^\#) \to \mathcal{B}_3/\Delta^2$ to $\boldsymbol{SL}(2, \mathbb{Z})$, see 3.3.8, and a choice of one of the two opposite real structures (which is the reason why the theorem is stated about pairs of opposite surfaces).

If the underlying surface of the ribbon curve structure is a disk, then the group $\pi_1(B^\#)$ is generated by small loops α_i around the singular fibers, the classes β_j of the hyperbolic components of $B_\mathbb{R}$, and the doubles γ_k of the dotted segments connecting the components of $B_\mathbb{R}$. The monodromy along each loop α_i and its lift to $\boldsymbol{SL}(2,\mathbb{Z})$ are determined by the requirement that the singular fibers of the surface must be of type I_1. The $\mathcal{B}_3/\Delta^2$-valued monodromy along each loop β_j or γ_k is trivial, see 5.3.5, and, hence, its lift to $\boldsymbol{SL}(2,\mathbb{Z})$ is $\pm\,\mathrm{id}$. The sign in front of id is the sign assigned to the corresponding dotted component/segment, and the collection of signs thus obtained (as well as the number of hyperbolic components) is encoded by the additional decoration in the definition of ribbon surface structure: the vertices subdividing the junctions into segments represent even dotted components, and a chain of $2m$ signs assigned to a box represents m even dotted components on the dotted segment inside the box attached to the oval. (Lemma 5.6.7 asserts that such a segment always exists.) In the latter case, the even numbered signs in the chain are those assigned to the dotted components (counted starting from the oval), and the odd numbered ones are the signs assigned to the segments connecting two consecutive components (or the first component and the oval).

The relation between the signs required in the definition of the ribbon surface structure is a manifestation of a relation between equivariant characteristic classes of a line bundle. In simple terms, it can be obtained as follows. As explained in 3.2.6 and 3.3.8, the homological invariants form an affine space over $\mathbf{H}^1(B;\mathbb{Z}_2)$, whence the signs involved are subject to one relation for each ribbon box. Pick a box, cut it out of the underlying surface, and cut out a disk containing the inner monochrome components. The result can be regarded as a cubic with a few half disks at the boundary removed. By an affine shift the signs can be chosen so that the monodromy along each cut but one is trivial. If it were also trivial on the remaining cut, it would extend to the whole cubic, thus producing an elliptic surface over a trigonal curve of odd degree, which is a contradiction.

If the underlying surface is an annulus, there is an additional pair of complex conjugate cycles, on which a lift should be chosen. This accounts for the fact that, in this case, the ribbon surface structure defines two pairs of deformation classes. □

Remark. If the underlying surface is an annulus, the two lifts of the monodromy along an additional cycle are also topologically distinct, see Lemma 3.2.3. Instead of choosing a lift, which depends on a particular choice of the cycle, one can distinguish the two surfaces by the orientability of the principal component over one of the two boundary components of the annulus, see 3.3.6.

6.3.2 COROLLARY. *The deformation classes of pairs of opposite almost generic elliptic M-surfaces over a rational base are in a canonical one-to-one correspondence with the isomorphism classes of ribbon curve structures with the following* *properties:*

- *each box represents an M-cubic, and the number of boxes is even;*
- *there are no vanishing segments;*
- *the junction graph is a linear tree.* □

6.4. ($M-1$)-curves and surfaces.

6.4.1 THEOREM. *The collection of all dotted cuts of any totally reduced nonhyperbolic* $(M-1)$*-dessin* Γ *represents* Γ *as an iterated (self-)junction of a union of* M*-cubics and at most one* $(M-1)$*-block, the latter being either a sextic* (i.e., a dessin of degree six on a disk) *or a block of degree* 3. *Furthermore, any elementary move in* Γ *is either a simple modification of the junction, or an elementary move in one of the blocks* (not involving the cuts separating distinct blocks)*, or an elementary move in a sextic that is the junction of the* $(M-1)$*-cubic and an* M*-cubic.*

Proof. The proof is almost literally the same as in the case of M-curves. Let $\Gamma \subset D$ be a totally reduced nonhyperbolic $(M-1)$-dessin that is neither of degree 3 nor a sextic. One has $d = 1$; hence, either

$\delta = 1$, $n_{\mathrm{z}} + n_{\mathrm{i}} = 0$, and D is a Möbius band, or

$\delta = 0$ and $n_{\mathrm{z}} + n_{\mathrm{i}} \leq 3$ if D is a disk, and $n_{\mathrm{z}} + n_{\mathrm{i}} \leq 2$ otherwise,

see 5.4. Thus, according to Theorem 5.7.6, the dessin Γ has a dotted cut. The oval count 5.4.1 implies that a dotted cut in Γ must be a (self-)junction, the result being an M- or $(M-1)$-dessin.

For the second statement, an elementary move not as in the theorem would destroy a junction and create a new one (due to the first statement). Hence, at the very moment of the modification the graph would have a genuine dotted cut that is not a junction. □

6.4.2 THEOREM. *Any totally reduced nonhyperbolic* $(M-1)$*-dessin is equivalent to a dessin that admits a ribbon curve structure such that either:*

- *each box represents an* M*-cubic and has valency* 2 (so that the junction graph is a single cycle)*, and the underlying surface is a Möbius band, or*
- *exactly one box represents an* $(M-1)$*-cubic, while all other boxes represent* M*-cubics, and the underlying surface is orientable.*

Conversely, any ribbon curve structure as above defines a totally reduced nonhyperbolic $(M-1)$*-dessin. Furthermore, two such dessins are equivalent if and only if their ribbon curve structures are isomorphic or connected by one or several of the following moves:* $\mathrm{II}_2 + \mathrm{I}_2 \leftrightarrow \mathrm{I}_2 + \mathrm{II}_2$ *provided that the adjacency is positive,* $\mathrm{II}_2 + \mathrm{I}_1 \leftrightarrow \mathrm{I}_2 + \mathrm{II}_1$, $\mathrm{II}_1 + \mathrm{I}_2 \leftrightarrow \mathrm{I}_1 + \mathrm{II}_2$, *or* $\mathrm{II}_1 + \mathrm{I}_1 \leftrightarrow \mathrm{I}_1 + \mathrm{II}_1$.

Proof. First, show the decomposability. In view of Theorem 6.4.1, it suffices to consider a dessin Γ that either is a sextic on a disk or has degree 3. Theorem 5.6.1 implies that Γ is equivalent to a dessin with a cut. It remains to consider all gluings of one or two cubics, select those that are $(M-1)$-curves, and, in each case, see that the dessin is equivalent to a junction, making sure that the equivalence

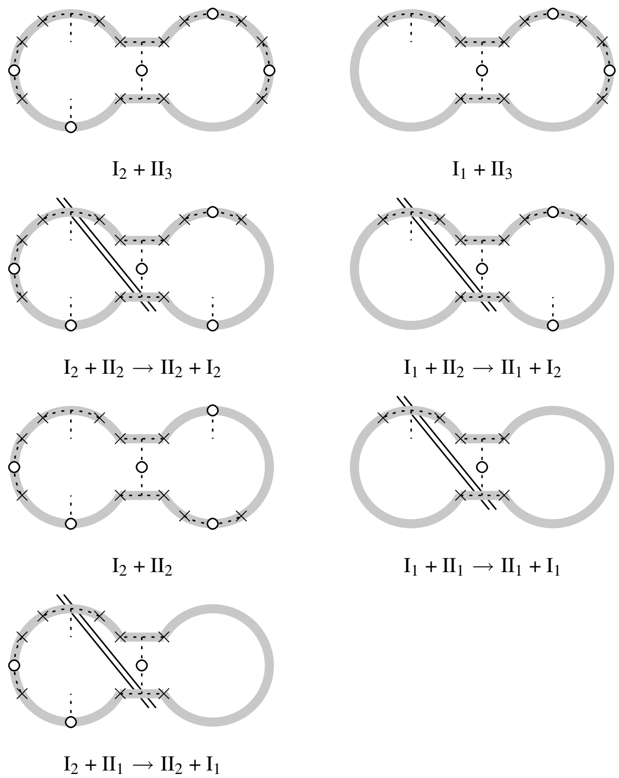


Figure 29. (Re-)decompositions of $(M - 1)$-sextics.

leaves intact the zigzags. This is straightforward. In fact, from Proposition 5.4.1 it follows that, since Γ is an $(M - 1)$-dessin, the cut is either a $\times$-cut (the result of the cut being an M-curve) or a dotted cut. A genuine dotted cut is necessarily a junction. An artificial dotted cut cannot join two real dotted segments adjacent to $\times$-vertices (as it must destroy two ovals); hence, creating a bridge, one can replace the cut with a genuine one.

The forms of the ribbon curve structures are easily enumerated using Proposition 5.4.1. According to Theorem 6.4.1, in order to study the equivalences, it suffices to consider an $(M - 1)$-sextic decomposed into a junction of two cubics, and study its re-decompositions. The sextics and their re-decompositions are shown schematically in Figure 29. (Recall that a box of type I_2 and a box of type II_2 can be joined in two different ways, forming a positive or negative junction.) Each re-decomposition shown (the double lines in the figure) can easily be realized by a sequence of elementary moves. A way to prohibit the other re-decompositions is to consider the distribution of the maximal real dotted/complementary segments and their parities. $\square$

Remark. The junction graph of an $(M - 1)$-dessin is either a linear tree, or a single cycle, or one of the two graphs shown in Figure 30, the vertex of

Figure 30. Extra junction graphs of $(M-1)$-dessins.

valency 3 representing the $(M-1)$-cubic, which is of type II_3. In the second graph in Figure 30, the cycle may as well consist of one or two vertices, cf. Figure 28.

6.4.3 Corollary. *The deformation classes of almost generic trigonal* $(M-1)$-*curves are in a canonical one-to-one correspondence with the isomorphism classes of enhanced ribbon curve structures as in Theorem 6.4.2 modulo the following additional equivalence relation:*

• *the moves as in Theorem 6.4.2; the three integers assigned to the two boxes and the adjacency involved can be chosen arbitrarily provided that the sum of the integers is left intact;*

• *the integers within a box of type* II_1 *or* II_2 (i.e., those assigned to the box itself, its adjacencies, and its vanishing segments) *can be changed arbitrarily provided that their sum is left intact.*

Proof. First, notice that an $(M-1)$-curve cannot be hyperbolic. Indeed, with $d = 1$ the oval count (5.4.2) implies $\delta = 0$, which contradicts Corollary 5.4.3. Thus, the curve is nonhyperbolic, and, similar to Theorem 6.2.3, the problem can be reduced to Theorem 6.4.2.

The realizability of both moves can be deduced from Lemma 5.6.7. As in the proof of Theorem 6.4.2, to show that there are no others, it suffices to consider sextic dessins. This can be done on a case by case basis, using Figure 29 and a careful analysis of real dotted monochrome vertices. □

6.4.4. In view of Corollary 4.3.7, the classification of almost generic elliptic $(M-1)$-surfaces also reduces to the classification of almost generic trigonal $(M-1)$-curves enhanced with a lift of the monodromy $\pi_1(B^\#) \to \mathcal{B}_3/\Delta^2$ to $\boldsymbol{SL}(2,\mathbb{Z})$. As in the case of M-surfaces, this procedure could be expressed in terms of ribbon surface structures. An additional complication is the fact that $(M-1)$-curves enjoy much more freedom (the equivalence relations described in Corollary 6.4.3) resulting in a vast number of moves for the monodromy. For this reason, we confine ourselves to the case of genus zero, where the monodromy is uniquely determined by the dessin and, hence, the result is an immediate consequence of Corollary 6.4.3.

6.4.5 Proposition. *The deformation classes of pairs of opposite almost generic elliptic* $(M-1)$-*surfaces over a rational base are in a canonical one-to-one correspondence with the isomorphism classes of ribbon curve structures with the fol-*

lowing properties:

- *the number of boxes is even;*
- *one box represents an* $(M-1)$*-cubic, the others representing* M*-cubics;*
- *there are no vanishing segments;*
- *the junction graph is a tree.* □

6.5. Oval chains. In this section we derive a few simple consequences of the classification results obtained above.

6.5.1 THEOREM. *Let* C *be a nonsingular trigonal* M*-curve of degree* $3k$ *on a real ruled surface over a base* B*. Then the following holds:*

(1) *each non-complete maximal chain of ovals of* C *is of odd length;*

(2) *if* C *has no complete chains, then it has* $k-2+\ell_{\mathrm{odd}}$ (respectively, k) *maximal chains if* $B_{\mathbb{R}}$ *has one* (respectively, two) *nonhyperbolic components.*

Proof. Corollary 6.2.2 lists all dessins of nonhyperbolic M-curves, and the maximal chains of ovals are easily seen: the ovals are described in 5.3.6, and chain breaks are the maximal sequences consisting of an odd number of odd segments (and any number of even segments other than ovals). □

6.5.2 COROLLARY. *The ovals of a nonsingular trigonal* M*-curve on a real rational ruled surface* Σ_k*,* $k \geq 3$*, form* $k-2$ *maximal chains, each maximal chain being of odd length.* □

Let $C' \subset \Sigma' \to B'$ and $C'' \subset \Sigma'' \to B''$ be two real trigonal curves on real ruled surfaces. The curves C' and C'' are said to have the same *fibered real scheme* if there is a fiberwise homeomorphism φ: $\Sigma''_{\mathbb{R}} \to \Sigma'_{\mathbb{R}}$ such that $C'_{\mathbb{R}}$ and the image $\varphi(C''_{\mathbb{R}})$ can be connected by an isotopy $C^t \subset \Sigma'_{\mathbb{R}}$ during which the intersection of C^t with any fiber of the projection $\Sigma'_{\mathbb{R}} \to B'_{\mathbb{R}}$ consists of at most three points. In other words, to make the result slightly more general, we allow passing through vertical flexes, i.e., straightening zigzags, cf. 6.6.2.

6.5.3 THEOREM. *There is a nonsingular trigonal* $(M-1)$*-curve on a rational ruled surface such that the fibered real scheme of the curve cannot be obtained by a single Morse modification from the fibered real scheme of a nonsingular* M*-curve.*

Proof. In fact, any $(M-1)$-curve whose junction graph has a vertex of valency three and three branches of length at least two each has the desired property. An example is shown in Figure 31. The reason is that the curve in question has three maximal chains of ovals of even length; hence, in view of Corollary 6.5.2, its fibered real scheme cannot be obtained by erasing one oval from the fibered real scheme of an M-curve. □

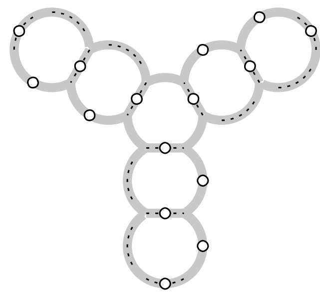


Figure 31. $(M-1)$-curve with three maximal chains of length six.

6.6. Further generalizations and open questions.

6.6.1. Singular curves. According to the definition given in 3.3.1, a trigonal curve in a ruled surface Σ is not supposed to intersect the distinguished section s of Σ. Relax this requirement and consider a curve $C \subset \Sigma$ that does intersect s at a point P. The elementary transformation of Σ at P (i.e., blowing up P and blowing down the fiber through P) produces a new surface Σ', section s', and curve $C' \subset \Sigma'$ which intersects s' with a smaller multiplicity; the image of the fiber blown down is a singular point of C' (a node or a cusp if C was nonsingular). Iterating this procedure, one arrives at a surface Σ'', section s'', and curve $C'' \subset \Sigma''$ disjoint from s'', i.e., a trigonal curve in the sense of 3.3.1. The curve is singular: it has one type $\mathbf{A}_{2m-1}$ or $\mathbf{A}_{2m}$ singular point for each m-fold intersection point of the original curve C and section s. The inverse elementary transformations (blowing up the singular points and contracting the corresponding fibers) convert C'' back to C. Thus, the deformation classification of trigonal (in the wide sense) curves intersecting s at several points with prescribed multiplicities can be reduced to that of trigonal (in the sense of 3.3.1) curves with several type $\mathbf{A}$ singular points. (Note that the degenerations $\mathbf{A}_{2m-1} \to \mathbf{A}_{2m}$ should be allowed during the deformations; these degenerations correspond to the confluences of vertical tangents and intersections with the exceptional section.) If the multiplicities are not prescribed, one should consider curves with a certain number of nodes and allow deeper confluence of the nodes during the deformations.

The classification of nonhyperbolic singular curves (with type $\mathbf{A}$ singularities only) that perturb to nonsingular M- or $(M-1)$-curves is essentially contained in Theorem 6.2.3 and Corollary 6.4.3. Indeed, type $\mathbf{A}$ singularities are obtained by bringing together some of the vertical tangents of a nonsingular curve. Hence, the graph (of the j-invariant) of the singular curve is obtained from that of a nonsingular one by bringing together some of the $\times$-vertices. Obviously, several consecutive $\times$-vertices can be brought together if and only if, after a sequence of $\circ$-ins and $\bullet$-ins, they are not separated by $\circ$- or $\bullet$-vertices. This observation gives a clear description of the singular curves in question, by either referring to the ribbon curve structures of nonsingular curves and indicating the sequences

Figure 32. Straightening a zigzag.

of $\times$-vertices to be brought together, or else constructing singular curves directly from boxes of more general form (including the graphs of singular cubic) via a more general junction operation (allowing forming singular points instead of ovals).

6.6.2. Straightening zigzags. On the account of the principal tool used in this paper (constructing deformations of curves via deformations of j-invariants), we state our results in the language of equivariant fiberwise deformations without confluence of singular fibers. However, from the point of view of geometry of nonsingular curves, vertical flex should not be considered a singularity. Passing through a vertical flex during a deformation results in the removing (straightening) or creating a zigzag. In spite of its apparent simplicity, this operation does not lead to a deformation of the j-invariant: at the very moment of the modification the degree of j drops by 2. The corresponding modification of dessins is shown in Figure 32 (see also [Z]): two adjacent triangles are removed, and two new triangles are inserted. Note that forming a vertical flex is not as local as forming a type **A** singular point (cf. 6.6.1); the possibility to bring together a pair of $\times$-vertices bounding a zigzag cannot be deduced solely from the real part. An example of a nonsingular trigonal curve with a zigzag that cannot be straightened in a single step was found by Orevkov [Or2].

Due to Theorems 6.2.1 and 6.4.1, whenever an M- or an $(M-1)$-curve is decomposed in ribbon boxes, each zigzag is localized within a single cubic, and using Lemma 5.6.7 one can conclude that each zigzag can be straightened without destroying the junctions. Hence, the classification of nonsingular M- and $(M-1)$-curves up to the new relaxed equivalence relation can be deduced from Theorems 6.2.1 and 6.4.1. We refrain from attempting to formulate a precise statement for $(M-1)$-curves. Just note that straightening a zigzag increases the number of moves as in Theorem 6.4.1 that can be applied to the dessin. Effectively, this increased flexibility means that, after its disappearance, a zigzag can freely slide along a chain of ovals and reappear at a new place. In the M-case, the ribbon curve structure is always rigid. Thus, the only modification is the disappearance of a zigzag, possibly followed by its reappearance next to the same oval, pointing to it from the other side. At the level of the ribbon curve structure, this modification is either a change of type (I_1 to I_2 or *vice versa*) of the corresponding cubic (which is necessarily located at one of the two ends of the junction graph) or a change of the sign of its junction.

6.6.3. Ribbon vs. unstructured curves. The ribbon curve construction produces an interesting class of trigonal curves with a clearly defined structure. Under various mild assumptions (e.g., if all $\times$-vertices are real, i.e., assuming that all boxes are of type I_2 or II_3) the structure is rigid: the junctions are present in any dessin equivalent to a given one, and they cannot be destroyed or modified by elementary transformations. Thus, the correspondence between curves and ribbon curve structures gives a deformation classification of such curves. On the other hand, there obviously are large "unstructured" curves whose graphs do not contain a junction or even a cut. At present, it is unclear whether and how the property of being a ribbon curve can be characterized in topological terms, or what a general classification theorem would look like. Probably, the dessin of a general curve would be a union of a ribbon part and a few unstructured pieces.

The simplest example of an "unstructured" curve is found among $(M-2)$-sextic. Indeed, one can glue two cubics of type I_2 each along a pair of solid segments so that the resulting sextic has two ovals not separated by zigzags (and hence four zigzags not separated by ovals). On the other hand, any sextic that is a junction of two cubics may have at most two zigzags within each real segment connecting the two ovals produced by the junction.

This example shows that, starting from the $(M-2)$-case, the structure theorems should have a form different from the classification results of the paper. (That is why we confine ourselves to M- and $(M-1)$-curves only.) Both of the key ingredients used in our approach fail: first, Theorem 5.7.6 does not break a dessin into sufficiently small pieces; second, it is no longer true that any dotted cut is a junction (note that the sextic above is equivalent to a dessin with a dotted cut).

6.6.4. Quasi-simplicity: still open. Let us briefly discuss the relation between the ribbon box decomposition of a trigonal M-curve (without hyperbolic components) and its real part. Certainly, the former does determine the latter, and there is a number of situations in which the converse also holds, i.e., the ribbon curve structure is recovered from the sequence formed by the zigzags and maximal chains of ovals. Among these situations are:

(1) M-curves over a base B of genus one;

(2) M-curves with at least one zigzag;

(3) M-curves with a sufficiently generic (in the sense described below) distribution of ovals.

In case (1) the sequence of maximal chains of ovals in one of the two components of $B_{\mathbb{R}}$ clearly determines the box decomposition. In the case of rational base the junction graph is a linear tree and there are two distinguished ovals which are located in the cubics corresponding to its two extreme vertices; we call them *extreme* ovals. If at least one extreme oval is known, the rest of the ribbon curve structure is found uniquely. This observation covers case (2), as zigzags of an M-curve always point at its extreme ovals.

In fact, the ribbon curve structure can still be recovered, at worst up to the "horizontal" symmetry, starting from one extreme chain, i.e., chain containing an extreme oval. If the base is rational and there are no zigzags, each maximal chain of length $2k+1$ containing $l = 0$, 1, or 2 extreme ovals is opposed by a maximal sequence of $k-l$ solitary ovals (i.e., those forming maximal chains of length 1). Hence, the lengths of the extreme chains are recovered from the sequence of maximal chains. By case (3) we mean the situation when these lengths determine the extreme chains.

However, there are sequences of maximal chains that can be obtained from two nonisomorphic ribbon curve structures. The simplest example that we know is the following sequence of 24 chains:

$$\underline{5}\ 1\ 3\ 3\ 1\ 3\ 3\ 5\ \underline{\underline{5}}\ 3\ 1\ 3\ \underline{5}\ 3\ 5\ 3\ 1\ 3\ 1\ 3\ 3\ \underline{\underline{5}}\ 3\ 5\ .$$

(Of course, only the lengths of the chains are listed.) The chains containing extreme ovals are either those underlined, or those double underlined. As the sequence have no symmetries, the two ribbon curve structures are not isomorphic.

BILKENT UNIVERSITY, 06800 ANKARA, TURKEY
E-mail: degt@fen.bilkent.edu.tr

UNIVERSITÉ LOUIS PASTEUR ET IRMA (CNRS), 7 RUE RENÉ DESCARTES 67084 STRASBOURG CEDEX, FRANCE
E-mail: itenberg@math.u-strasbg.fr

UNIVERSITÉ LOUIS PASTEUR ET IRMA (CNRS), 7 RUE RENÉ DESCARTES 67084 STRASBOURG CEDEX, FRANCE
E-mail: kharlam@math.u-strasbg.fr

REFERENCES

[AMn] M. Akriche and F. Mangolte, Nombres de Betti des surfaces elliptiques réelles, preprint ArXiv: math.AG/0604131, 2006.

[Ba] E. Ballico, Real analytic elliptic surfaces, *J. Pure Appl. Algebra* **178** (2003) no. 3, 225–234.

[BPV] W. Barth, C. Peters, and A. van de Ven, *Compact Complex Surfaces*, *Ergeb. Math. Grenzgeb.*, vol. 3, Springer-Verlag, Berlin-New York, 1984.

[BE] I. Bernstein and A. L. Edmonds, On the classification of generic branched coverings of surfaces, *Illinois J. Math.* **28**(1) (1984).

[BMn] F. Bihan and F. Mangolte, Topological types of real regular jacobian elliptic surfaces, *Geom. Dedicata*, **127** (2007), 57–73.

[Br1] G. E. Bredon, *Introduction to Compact Transformation Groups*, Academic Press, New York, London, 1972.

[Br2] ———, *Sheaf Theory*, 2nd edn., *Grad. Texts in Math.*, 1997, Springer-Verlag, New York, Berlin, Heidelberg.

[Bro] K.S. Brown, *Cohomology of Groups*, Springer-Verlag, New York-Heidelberg-Berlin, 1982.

[CF] F. Catanese and P. Frediani, Real hyperelliptic surfaces and the orbifold fundamental group, *J. Inst. Math. Jussieu* **2**(2) (2003), 163–233.

[Co1] A. Comessatti, Fondamenti per la geometria sopra le superficie razionali dal punto di vista reale, *Math. Ann.* **43** (1912), 1–72.

[Co2] ———, Sulle varietà abeliane reali, I e II, *Ann. Mat. Pura Appl.* **2**, **4** (1924, 1926), 67–106, 27–71.

[DIK1] A. Degtyarev, I. Itenberg, and V. Kharlamov, *Real Enriques Surfaces*, *Lecture Notes in Math.*, vol. 1746, 2000, Springer-Verlag.

[DIK2] ———, Finiteness and quasi-simplicity for symmetric K3-surfaces, *Duke Math. J.* **122** (2004), 1–49.

[DK1] A. Degtyarev and V. Kharlamov, Real Enriques surfaces without real points and Enriques-Einstein-Hitchin 4-manifolds, *The Arnoldfest (Toronto, ON, 1997)*, *Fields Inst. Commun.*, vol. 24, Amer. Math. Soc., Providence, RI, 1999, pp. 131–140.

[DK2] ———, Real rational surfaces are quasi-simple, *J. Reine Angew. Math.* **551** (2002), 87–99.

[Fr] P. Frediani, Real Kodaira surfaces, *Collect. Math.* **55** (2004) no. 1, 61–96.

[FM] R. Friedman and J. W. Morgan, *Smooth Four-Manifolds and Complex Surfaces*, *Ergeb. Math. Grenzgeb.*, vol. 3, Springer-Verlag, Berlin-New York, 1994.

[GW] M. Gross and P. M. H. Wilson, Mirror symmetry via 3-tori for a class of Calabi-Yau threefolds, *Math. Ann.* **309**(3) (1997), 505–531.

[Ka] A. Kas, On the deformation types of regular elliptic surfaces, *Complex Analysis and Algebraic Geometry*, Iwanami Shoten, Tokyo, 1977, pp. 107–111.

[Kh] V. Kharlamov, Towards the Viro, Ragsdale, and Petrovskii conjectures, *Proc. Internat. Topological Conference, Baku, 1987*, p. 412.

[KK] V. Kharlamov and Vik. Kulikov, Surfaces with DIF $\neq$ DEF real structures, *Izv. Math.* **70** (2006), no. 4, 769–807.

[Kl] F. Klein, Ueber Flächen dritter Ordnung, *Math. Ann.* **6** (1873), 551–581.

[Ko] K. Kodaira, On compact analytic surfaces, II–III, *Ann. of Math.* **77–78** (1963), 563–626, 1–40.

[Mn] F. Mangolte, Surfaces elliptiques réelles et inégalité de Ragsdale-Viro, *Math. Z.* **235** (2000) no. 2, 213–226.

[Ma] Yu. Manin, Rational surfaces over perfect fields, *Mat. Sbornik* (1967) no. 1, 141–168.

[N1] S. Natanzon, Klein surfaces, *Uspekhi Mat. Nauk* **45** (1990), 47–90.

[N2] ———, Topology of 2-dimensional covering and meromorphic functions on real and complex algebraic curves, *Selecta Math.* **12**(1) (1993), 251–291.

[NSV] S. Natanzon, B. Shapiro, and A. Vainshtein, Topological classification of generic real rational functions, *J. Knot Theory Ramifications* **11** (2002) no. 7, 1063–1075.

[Ni] V. Nikulin, Integer quadratic forms and some of their geometrical applications, *Math. USSR–Izv.* **43** (1979), 103–167.

[Or1] S. Orevkov, Riemann existence theorem and construction of real algebraic curves, *Ann. Fac. Sci. Toulouse Math.* (6) **12**(4) (2003), 517–531.

[Or2] ———, Private communications.

[P] A. Protopopov, Topological classification of branched coverings of the two-dimensional sphere, *Zapiski Nauchnyh Seminarov LOMI*, vol. 167, 1988, pp. 135–156.

[Se] W. Seiler, Moduln elliptischer Flächen mit Schnitt, Karlsruhe, Thesis, 1982.

[SV] B. Shapiro and A. Vainshtein, Counting real rational functions with all real critical values, *Moscow Math. J.* **3** (2003) no. 2, 647–659.

[Si] R. Silhol, *Real Algebraic Surfaces*, *Lecture Notes in Math.*, vol. 1392, Springer-Verlag, 1989.

[Wa] C. T. C. Wall, Real forms of smooth del Pezzo surfaces, *J. Reine Angew. Math.* **375/376** (1987), 47–66.

[We] J.-Y. Welschinger, Real structures on minimal ruled surfaces, *Comment. Math. Helv.* **78**(2) (2003), 418–446.

[Z] V. Zvonilov, Rigid isotopies of threenomial curves with the maximal number of ovals, *Vestnik Syktyvkarskogo Universiteta* **6** (2006), 23 pages.