REMARKS ON THE PARABOLIC CURVES ON SURFACES AND ON THE MANY-DIMENSIONAL MÖBIUS–STURM THEORY

V. I. Arnold

V. A. Steklov Mathematical Institute

August 26, 1997

Conjecture on Four Parabolic Curves

According to the classical Möbius theorem, a curve on the projective plane has at least three inflection points provided that this curve is close to a projective line. This theorem is a special case of the Sturm–Hurwitz theorem [2] stating that the number of the sign alternations of a function on a circle is not less than the number of the sign alternation of the lowest harmonic entering its Fourier series with nonzero coefficient. The Möbius theorem can be extended to all noncontractible curves embedded in the projective plane. The Sturm theorem has probably a global version too. Attempts to extend the Möbius and Sturm theorems to the many-dimensional case (i.e., to surfaces in the projective space and to functions of several variables) show that natural analogs of inflection points are parabolic curves of surfaces and zeros of the Hessians of functions. This means that the differential operators corresponding to surfaces and to functions of two variables are quadratic (of degree n in the case of n-dimensional manifolds and functions of n variables) rather than linear (as is the case in the ordinary Sturm theory of functions of one variable). According to the general principle of topological economy of algebro-geometric objects, one might expect that the minimal number of parabolic curves (say, on surfaces which are small deformations of a projective plane in the projective space) is attained in the case of cubic surfaces. For such cubic surfaces, the number of parabolic curves (calculated by B. Segre [1]) is four. This leads to the *Aicardi conjecture* (see [3]): *the number of parabolic curves on a generic surface in the projective space is at least four provided that the surface is sufficiently close to a projective plane* (*in the smooth topology*). One might even conjecture (by analogy with the Möbius theorem and with the theory of Lagrangian intersections and Legendre linkings in symplectic and contact topology) that the number of parabolic curves is still not less than four for large deformations of the projective plane as well (at least while the deformed surface remains embedded) and that the corresponding results of the "set-valued Morse theory

---

Supported by the Russian Foundation for Basic Research, project 96-01-01104.

for higher derivatives" hold for more general problems of symplectic and contact topology with nonholonomic constraints. However, at present even the simplest special case of these conjectures stated above is neither proved nor disproved. This special case will be studied below. Our strategy is as follows. We choose a construction of perturbations of the projective plane for which we can inspect the number of parabolic curves. We try to choose the construction parameters so that the deformed surface would have less than four parabolic curves. If such a choice is possible, then we find a counterexample to the conjecture. If, for any choice of the parameters, we obtain at least four parabolic curves, then we receive a confirmation of the conjecture and some (possibly nontrivial) information on the deformations of the class under consideration (for instance, on the "spherical functions with singularities" introduced below). Possibly the most natural approach is to consider the polyhedral deformations constructed from the triangulations of the perturbed projective plane. On deforming the resulting polyhedron into a generic smooth surface by some specially chosen construction we might probably describe the number of occurring parabolic curves on the surface in terms of the combinatorics of the polyhedron (taking into account the singular curvatures concentrated at edges and vertices). In this way one might reformulate topological conjectures on parabolic curves as combinatorial statements on the triangulations of the projective plane. However, in the present paper, such a smoothing is not introduced and, instead of a combinatorial perturbation, we consider another perturbation of the projective plane. This perturbation is hyperbolic almost everywhere except for finitely many points including the logarithmic poles (in a neighborhood of each of this poles, the perturbation can be smoothed on creating a small parabolic curve and an elliptic component). These "hyperbolic" perturbations of the projective plane are constructed by means of the singular odd spherical functions u (that satisfy the equation $\Delta_S u + 2u = g$ on S^2, where Δ_S is the spherical Laplacian and g is a linear combination of delta functions). We shall prove that *the number of parabolic curves on a deformed surface generated by this construction is at least four.* Another construction of perturbations of the projective plane involves the decomposition of the projective plane into a (large) disk and a (thin) Möbius band, and a perturbation is performed on one of this two domains. For instance, one can treat a perturbation defined by a cubic surface over the Möbius band. Now, to study the perturbations over the disk, it suffices to study the zeros of the Hessian of the function that defines the perturbation in an affine coordinate system and satisfies given boundary conditions (provided by the chosen perturbation outside the disk). Thus, we arrive at the problem on the plane parabolic curves defined by the equation $F_{xx}F_{yy} = F_{xy}^2$. The study of these curves is equivalent to the study of the critical points and caustics of the gradient mappings. Our results imply, for instance, lower estimates for the number of cusps on the caustics of a gradient mapping of the disk satisfying given boundary conditions. For example, *if a generic potential coincides with the function $F = x^3 - 3xy^2$ outside some large disk, then the gradient mapping has at least three A_3 singularities inside this disk (which means that the caustic has at least three cusps).* This example shows that the gradient mappings are distinguished by their topological properties in the larger class of Lagrangian mappings (in which our mapping can be deformed into a mapping with only one cusp, according to Entov and Eliashberg). Another way to inspect parabolic curves is to start from a surface with a degenerate parabolic curve and to deform it into a generic surface (in a neighborhood of the curve). For instance, for the surface $z = 1/(x^2 + y^2)$, the line at infinity is a degenerate parabolic curve. In this case and in several similar examples we show that, *under a generic one-parameter small deformation of the surface, a degenerate parabolic curve decomposes into nondegenerate parabolic curves whose total number is at least four.* In the above example, the line at infinity gives birth to at least three parabolic components (actually, by the Möbius–Sturm theorem). I conjecture that, in fact, the results that are proved below only for small one-parameter deformations are valid in an entire neighborhood of our surface in the space of deformed surfaces (not only in our special examples but also for all surfaces with degenerate (multiple) parabolic curves). Should the Aicardi conjecture be true, it would prove these results even for very large deformations of the surface provided that this surface remains fixed outside a neighborhood of the degenerate parabolic curve.

§2. Equations of Parabolic Curves

We describe the projective space $\mathbb{R}P^3$ by means of the homogeneous coordinates $[x\!:\!y\!:\!z\!:\!w]$. We start from the unperturbed plane $\mathbb{R}P^2$ defined by the equation $w = 0$. Any perturbed surface is defined by an equation of the form where f is a homogeneous function of degree one (and thus an odd function). Let us find the parabolic curves on the perturbed surface. The following result is classical.

Lemma 1. *The parabolic points of the surface $w = f$ are represented on the sphere $x^2 + y^2 + z^2 = 1$ by the critical points of the restriction of the gradient of the function f to this sphere.*

Proof. Near a parabolic point, a surface is exceptionally well approximated by a plane. A plane in $\mathbb{R}P^3$ is defined by a linear function f. The gradient of linear function is constant. If f is exceptionally well approximated by a linear function at a given point, its gradient is exceptionally well approximated by a constant. This means that this point is critical for the gradient of f. □

Here one can replace the sphere (in a neighborhood of the point under consideration) by any surface transversal to the radial rays (since the gradient of a homogeneous function of degree one is constant along each of these rays). For instance, one may take the plane tangent to the sphere at this point.

Example. The domain $z \neq 0$ is served by the plane $z = 1$. Denote the restriction of f to this plane by $F(x, y) = f(x, y, 1)$. We may consider the relation $w = F(x, y)$ as the affine equation of the perturbed surface in the affine chart $z = 1$ of the projective space. We obtain the following obvious formulas.

Lemma 2. *In our affine coordinates, the equation of the curve of parabolic points of the surface $w = f$ acquires the form*

$$\operatorname{Hess} F = \det \begin{pmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{pmatrix} = 0. \tag{1}$$

Proof. For $z = 1$, the equation of the tangent plane $w = ax + by + cz$ becomes $w = ax + by + c$. Hence, the coefficients a, b, and c are equal to the values of F_x, F_y, and F, respectively, at our point. The deviation of the surface from its tangent plane is measured (in the quadratic approximation) by the second differential of F at the tangency point. This quadratic form is degenerate if and only if $\operatorname{Hess} F = 0$. □

Introduce the *polar coordinates* ($x = R\cos\varphi$, $y = R\sin\varphi$) on the plane $z = 1$. We shall use the archaic notation for the derivatives: $F_R = (\partial/\partial R)F(R\cos\varphi, R\sin\varphi)$, and so on.

Lemma 3. *In the polar coordinates, the equation for the curve of parabolic points of the surface $w = f$ reads*

$$\det \begin{pmatrix} F_{\varphi\varphi} + RF_R & F_{\varphi R} - F_\varphi/R \\ F_{\varphi R} - F_\varphi/R & F_{RR} \end{pmatrix} = 0. \tag{2}$$

Proof. Choose a point (R_0, φ_0). The tangent plane at this point has the equation $w = aR\cos\varphi + bR\sin\varphi + c$. By equating the 1-jet of F with the 1-jet of the right-hand side of this equation at our point, we obtain the following equations for (a, b, c):

$$\begin{aligned} F^{(0)} &= aR_0\cos\varphi_0 + bR_0\sin\varphi_0 + c, \\ F_R^{(0)} &= a\cos\varphi_0 + b\sin\varphi_0, \\ F_\varphi^{(0)} &= -aR_0\sin\varphi_0 + bR_0\cos\varphi_0 \end{aligned}$$

(the superscript 0 stands for the evaluation at our point). The second derivatives of the difference $F - aR\cos\varphi - bR\sin\varphi - c$ at our point form the matrix on the left-hand side in (2). Its determinant vanishes just at the parabolic points.

To study the surface in a neighborhood of the point at infinity ($z = 0$), it is convenient to use the *cylindrical coordinates* ($r = 1/R$, φ). This chart of the projective plane is defined by the central projection of the sphere $x^2 + y^2 + z^2 = 1$ on the cylinder ($x = \cos\varphi$, $y = \sin\varphi$, $z = r$) that is tangent to the sphere along the equator. The opposite points of the cylinder, (r, φ) and $(-r, \varphi + \pi)$, represent the same point of the projective plane. The parabolic curve is cooriented, and hence each of its components is represented on the cylinder by two ovals. A point ($x = R\cos\varphi$, $y = R\sin\varphi$, $z = 1$) of the plane $z = 1$ corresponds to the point ($x = \cos\varphi$, $y = \sin\varphi$, $z = r = 1/R$) of the cylinder. On this cylinder, the homogeneous equation of the perturbed surface $w = f(x, y, z)$

becomes $w = R^{-1}F(R,\varphi)$ (certainly, we mean the expression of the function F defined on the plane via the polar coordinates). Let us rewrite the last equation in the form $w = H(r,\varphi)$. The functions H and F are related as follows:

$$H(r,\varphi) = rF(1/r,\varphi), \qquad F(R,\varphi) = RH(1/R,\varphi). \tag{3}$$

The function H is smooth at $r = 0$. It satisfies the relation

$$H(-r,\varphi+\pi) = -H(r,\varphi)$$

because the function $w = f(x,y,z)$ is odd. Hence, the Taylor coefficients of the expansions

$$H(r,\varphi) = h_0(\varphi) + rh_1(\varphi) + \frac{r^2}{2}h_2(\varphi) + \dots,$$
$$F(R,\varphi) = Rh_0(\varphi) + h_1(\varphi) + \frac{1}{2R}h_2(\varphi) + \dots$$

satisfy the oddness conditions

$$h_{2k}(\varphi+\pi) = -h_{2k}(\varphi), \qquad h_{2k+1}(\varphi+\pi) = h_{2k+1}(\varphi). \tag{4}$$

Lemma 4. *In the cylindrical coordinates, the equation of the parabolic curve of the surface $w = f$ has the following form at infinity:*

$$\det\begin{pmatrix} H_{\varphi\varphi} + H - rH_r & H_{r\varphi} \\ H_{r\varphi} & H_{rr} \end{pmatrix} = 0. \tag{5}$$

Proof. Arguing as in the proof of Lemma 3, we obtain $w = a\cos\varphi + b\sin\varphi + cr$. Hence, (a,b,c) satisfies the equations

$$H^{(0)} = a\cos\varphi_0 + b\sin\varphi_0 + cr_0, \quad H_r^{(0)} = c, \quad H_\varphi^{(0)} = -a\sin\varphi_0 + b\cos\varphi_0.$$

The second derivatives of the difference $H - a\cos\varphi - b\sin\varphi - cr$ form the matrix in Eq. (5). $\square$

We also need the following assertion.

Lemma 5. *Suppose that a homogeneous function f of degree one is harmonic. Then the corresponding quadratic form that measures the deviation of the surface $w = f$ from its tangent plane is either hyperbolic or zero (thus, the above three Hessians (1), (2), (5) are either negative or zero; in the latter case, all entries of the corresponding matrix vanish).*

Proof. Without loss of generality we can assume that our point has the coordinates $(x = y = 0,\ z = 1)$. At this point we have $f_{zz} = 0$ because f is homogeneous of degree one; however, $f_{xx} = F_{xx}$ and $f_{yy} = F_{yy}$. Hence, the trace $F_{xx} + F_{yy}$ of the Hessian matrix is zero. Thus, at our point we have $\operatorname{Hess} F = -F_{xx}^2 - F_{xy}^2 \le 0$. $\square$

Remark. The restriction of a first-order homogeneous harmonic function in the space to the unit sphere is an eigenfunction of the spherical Laplacian with eigenvalue -2, i.e., $\Delta_S u + 2u = 0$. The spherical functions of degree one are simply the linear functions $ax + by + cz$. However, we can also apply Lemma 5 to "singular spherical functions" satisfying $\Delta_S u + 2u = g$, where g is an odd distribution on the sphere with support consisting of finitely many points. The solutions of this equation define singular perturbations of the projective plane related to surfaces whose second quadratic form is hyperbolic everywhere except for the poles at the points of the support of g and the points of total degeneration (at which the entire form is zero). One can use these special singular perturbations to construct regular perturbations for which we can inspect their parabolic curves (with a desired curve near every logarithmic pole of u and near each point of total degeneration provided that g is a generic combination of delta functions).

Definition. A function on a sphere is said to be *hyperbolic* if the corresponding quadratic form is hyperbolic everywhere except for finitely many points (where the function may be singular). The simplest singularity of a hyperbolic function is a (logarithmic) pole.

Conjecture. *An odd hyperbolic function on the* 2*-sphere all of whose singularities are the simplest ones has at least six logarithmic poles.*

§3. Singular Spherical Functions

By a *singular spherical function* we mean a function satisfying the equation of spherical functions everywhere except for finitely many singular points (at which the result of application of the corresponding operator to the function is a finite-order distribution, that is, a linear combination of delta functions and their derivatives).

Lemma 6. *An odd singular spherical function of degree one on the* 2*-sphere all of whose singularities are the simplest* (*logarithmic*) *poles has at least six such poles. If it has exactly six poles, then they all belong to the same great circle of the sphere.*

Proof. By assumption, the function $\Delta_S u + 2u = g$ is a linear combination of finitely many delta functions with positive coefficients at k points and with opposite coefficients at the opposite points of the sphere. By the Fredholm alternative, the above equation for u is solvable if and only if g is orthogonal to the solutions of the equation $\Delta_S u + 2u = 0$, that is, to the linear functions. This orthogonality condition means that the mass center of the k positive charges of the distribution g coincides with the center of the sphere. The mass center of two points on the sphere coincides with the center only if the points are opposite and the masses are equal. In this case $g = 0$. Hence, $k \ge 3$. The mass center of three points on the sphere coincides with the center only if the plane containing these points passes through the center. In this case, the three opposite points belong to the same plane. $\square$

Theorem 1. *A generic surface obtained from the projective plane by a deformation generated by a generic odd singular spherical function with the simplest poles (by smoothing and by a small generic deformation of this function) has at least four parabolic curves.*

Proof. Near each of the poles, smoothing creates a small parabolic curve bounding an elliptic domain. The total vanishing of the form is impossible for the generic surfaces, but it can stably occur at individual points on hyperbolic surfaces. One can construct a hyperbolic surface by using spherical functions or solutions of other elliptic differential equations. For instance, for the usual harmonic functions $u = \operatorname{Re} U$ on the plane of the variable $z = x + iy$, the three second partial derivatives of u vanish at the zeros of the second derivative of the holomorphic function U, and these zeros persist under deformations of the holomorphic function. According to Lemma 6, the number of poles (and hence the number of parabolic curves generated by the poles under smoothing) is at least six. Besides these singularities, points of the total degeneration of the second differential must also exist. Let us prove this fact. *There is a field of asymptotic directions in the hyperbolic domain of the sphere (that is, in the complement of the poles and the points of total degeneration).* Indeed, a field of hyperbolic quadratic forms on the tangent planes of the sphere is well defined in the hyperbolic domain. A straight line bounding from the right the sector in which the form is positive is defined by the orientation of the sphere unambiguously. These (nonoriented) lines form the desired field of asymptotic directions. The sum of the indices of the singular points of this field is 2 (the indices can be half-integers because the lines of the field are not oriented). But the $2k$ poles contribute $2k$ to the index (since the index of the field of asymptotic lines is equal to one at every simplest pole (and possibly at any point at which the hyperbolic function tends to ∞)). Hence, the sum of indices of the points of total degeneration is $2 - 2k \leq -4$. In particular, the index of at least one of these points is nonzero. We shall prove that, under a generic deformation of a hyperbolic surface (or function), such a point generates a small parabolic curve. Indeed, consider the "second Gauss" mapping of the plane (x, y) into the three-dimensional space of the quadratic forms $A\,dx^2 + 2B\,dx\,dy + C\,dy^2$ given by the formulas $A = F_{xx}$, $B = F_{xy}$, and $C = F_{yy}$ (locally, in a neighborhood of our point of total degeneration, which we can assume to have the coordinates $(x = 0,\ y = 0,\ z = 1)$). The degenerate forms constitute the cone $AC = B^2$. The parabolic points are the points of the preimage of this cone under the second Gauss mapping. Since F is hyperbolic, it follows that a punctured neighborhood of the point $x = y = 0$ at which the form is totally degenerate is sent by the second Gauss mapping into the domain $AC < B^2$ of hyperbolic forms. This domain is the nonsimply connected exterior component of the complement of the cone. *The index of our isolated singular point of the field of asymptotic directions is proportional (with a nonzero coefficient) to the linking number of the image of the boundary circle of our neighborhood with the axis of the cone.* Hence, if the index is nonzero, then the linking coefficient is nonzero. After a small generic deformation of F (under which F becomes nonhyperbolic), the image of neighborhood no longer contains the vertex of the cone ($A = B = C = 0$), but this image intersects the cone along a closed curve. Hence, our point of total degeneration gives birth to a small parabolic curve under a generic deformation of the initial hyperbolic function F. Together with the three parabolic curves born near the three poles, we obtain at least four parabolic curves on $\mathbb{R}P^2$. $\square$

Remark. Our reasoning gives $3k-2$ parabolic curves on $\mathbb{R}P^2$ provided that the initial odd hyperbolic function on the sphere has $2k$ simplest poles and only the simplest points of total degeneration (which possibly holds for odd singular spherical functions under a generic distribution of the charges of g). Indeed, the simplest points of total degeneration are the monkey saddles similar to that defined by the harmonic function $F = \mathrm{Re}(x+iy)^3$ at the origin. In this simplest case, the index of the field of asymptotic directions is equal to $-1/2$, and the coefficient of linking with the axis of the cone is equal to -1 (if we choose the natural orientation on the plane $A + C = 0$ of zero-trace forms). In this case, the second Gauss mapping of a neighborhood of the point of total degeneration into the hyperbolic domain of the space of forms (A, B, C) is transversal to the axis of the cone at the point of total degeneration $x = y = 0$.

Hence, on the sphere we obtain N simplest points of total degeneration with the sum of the indices $2 - 2k = -N/2$. Hence, the number $N/2$ of the simplest points of total degeneration on $\mathbb{R}P^2$ is $2k-2$. Together with k poles, this gives $3k-2$ parabolic curves. If the points of total degeneration are not the simplest ones, then we seemingly can obtain only four parabolic curves, as can be shown by example in which equal positive charges form a regular triangle. One may obtain singular spherical functions explicitly from homogeneous rational harmonic functions in $\mathbb{R}^3$. On this way one can find the simplest odd singular spherical function of degree one whose only singularities are the poles of the sphere (the calculations are very simple in the coordinates $\xi = x + iy$, $\eta = x - iy$). These singularities are multipole: in our notation,

$$F = (2R^{-2} + 3)\cos 2\varphi.$$

Under a perturbation of the singular function F, one obtains at least four parabolic curves on $\mathbb{R}P^2$: two are born from the two (simplest) points of total degeneration at infinity and the other two at the maximum and at the minimum of the smoothed singularity at the origin.

Conjecture. *The number of parabolic curves generated by smoothing an odd hyperbolic function on the 2-sphere and under a small deformation making it generic, is at least four, for arbitrary finite-order poles.*

This is not proved even for the rational spherical functions of degree one (but this certainly follows from the Aicardi conjecture).

§4. Deformations of a Degenerate Parabolic Curve

Consider the cubic surface defined in the affine chart $z = 1$ by the equation $w = 1/(x^2 + y^2)$ (we shall later smooth the singularity at the origin). This cubic surface, for which $F = R^{-2}$ and $H = r^3$, is hyperbolic in its regularity domain over the (x, y)-plane and has a degenerate parabolic curve at infinity ($r = 0$). Perturbations of this surface near infinity are described in the cylindrical coordinates by the function $H = r^3 + \varepsilon h(r, \varphi)$, where $\varepsilon \ll 1$ and the perturbing function h satisfies the oddness condition $h(-r, \varphi + \pi) = -h(r, \varphi)$.

Theorem 2. *For a sufficiently small ε, the degenerate parabolic curve at infinity of the original surface transforms into at least three nondegenerate parabolic curves of the perturbed surface, provided that the perturbing function h is generic.*

Proof. The equation of parabolic points (5) has the form

$$-12r^4 + \varepsilon(6r(h_{\varphi\varphi} + h - rh_r) - 2r^3 h_{rr}) + O(\varepsilon^2) = 0. \tag{6}$$

Let us introduce a new variable ρ by the scaling $r = \varepsilon^{1/3}\rho$. Equation (6) becomes

$$\varepsilon^{4/3}[\rho(-12\rho^3 + 6a(\varphi)) + O(\varepsilon^{1/3})] = 0, \tag{7}$$

where $a = (h_{\varphi\varphi} + h)|_{r=0}$. First we neglect the term with O and obtain from (7) a curve consisting of two smooth branches, $\rho = 0$ and $\rho = (a(\varphi)/2)^{1/3}$, which intersect each other transversally at the zeros of the function a (which are simple for generic h). According to the Sturm theorem, *the number of these zeros is at least six.* Indeed, by (4), for $r = 0$ we have $h(\varphi + \pi) = -h(\varphi)$. Hence, all harmonics of h are of odd order. It follows that the lowest harmonic of a is of order not less than three and has at least 6 zeros. The product $\rho(-12\rho^3 + 6a(\varphi))$, where the zeros of a are simple, is a Morse function in neighborhoods of the saddle points placed at the $2k$ zeros of the function a. A perturbation of the product defined by the O-term in (7) moves these $2k$ saddles to nonzero levels, and thus the pair of intersecting branches at the zero level of the function (7) transforms into a nonsingular curve. The Hessian nature of Eq. (7) determines the signs of all $2k$ surgeries.

Lemma 7. *The smooth curve (7) defined by a generic perturbation h consists of $2k$ components that bound the domains close to the positivity domains of the product $\rho(-12\rho^3 + 6a(\varphi))$.*

Proof. Consider a plane in the space of quadratic forms $A\xi^2 + 2B\xi\eta + C\eta^2$ that intersects the cone of degenerate forms $AC = B^2$ along two lines. Under a small generic deformation of this plane we obtain a plane whose intersection with the cone forms a hyperbola, and the discriminant $AC - B^2$ is negative between the two branches of the hyperbola. In other words, from *the four quadrants of the nonperturbed plane that consist of elliptic and of hyperbolic forms alternately, the two opposite hyperbolic quadrants merge under a generic deformation, while the elliptic part becomes separated by the hyperbolicity domain on the perturbed plane.* Hence, the number of the components born at infinity is $2k \geq 6$ on the cylinder, that is, $k \geq 3$ on $\mathbb{R}P^2$. Theorem 2 is thereby proved. □

It follows from the theorem that smoothing our surface at the origin followed by a deformation making it generic at infinity makes it a smooth surface with at least four parabolic curves.

Remark. It seems that *the number of parabolic curves on any surface defined by a function of the form* $H = r^3 + h(r, \varphi)$*, where* $h(-r, \varphi + \pi) = -h(r, \varphi)$*, is at least three provided that the perturbation* h *is generic and sufficiently small (and this possibly holds for any generic* h *vanishing for* $|r| > 1$*).* It is proved above only for one-parameter families of deformations.

Now consider the cubic surface defined by the equation $w = x/(x^2 + y^2) = R^{-1} \cos \varphi$ (in the affine chart $z = 1$). The function is harmonic and hence hyperbolic (and singular at the origin). The surface has a degenerate parabolic curve at infinity (where the surface is defined by the function $H = r^2 \cos \varphi$ in the cylindrical coordinates of §2). Equation (5) of parabolic points becomes $-4r^2 = 0$. Near the line at infinity the perturbed surface is defined by the function

$$H = r^2 \cos \varphi + \varepsilon h(r, \varphi), \qquad \varepsilon \ll 1, \quad h(-r, \varphi + \pi) = -h(r, \varphi).$$

Theorem 3. *For sufficiently small values of the perturbation parameter* ε*, the degenerate parabolic curve at infinity is transformed into at least two nondegenerate parabolic curves of the corresponding perturbed surface provided that* h *is generic.*

We conjecture that it suffices to assume that the perturbation (which may not be small) vanishes for $|r| \geq 1$.

Proof. The reasoning in the proof of Theorem 2 leads (by (5)) to the equation

$$\text{(6')} \qquad -4r^2 + \varepsilon (h_{\varphi\varphi} + h)|_{r=0} 2 \cos \varphi + O(r\varepsilon) + O(\varepsilon^2) = 0$$

of the parabolic curve. By substituting $r = \varepsilon^{1/2} \rho$ we obtain

$$\text{(7')} \qquad \varepsilon(-4\rho^2 + 2a \cos \varphi + O(\varepsilon^{1/2})) = 0.$$

The product $a \cos \varphi$ (where $a = (h_{\varphi\varphi} + h)|_{r=0}$) has at least eight simple roots provided that h is generic (since a has at least six roots according to Sturm's theorem). The equation $-4\rho^2 + 2a(\varphi) \cos \varphi = 0$ defines at least four smooth ovals (corresponding to the intervals where $a \cos \varphi$ is positive) on the cylinder $\{(r, \varphi)\}$. The small perturbation $O(\varepsilon^{1/2})$ in (7') does not destroy them. On $\mathbb{R}P^2$ we obtain at least two parabolic ovals. □

Theorem 4. *Smoothing the function* $w = x/(x^2 + y^2)$ *at the origin followed by a generic small deformation of the resulting surface in the projective space defines a surface with at least 4 parabolic curves.*

Proof. Two parabolic curves appear at infinity according to Theorem 3. We shall prove that smoothing at the origin creates at least two another parabolic curves.

The proof is based on the existence of a maximum point and a minimum point of the smoothed function. Indeed, on completing the (x, y)-plane by one point and applying the stereographic projection we obtain the sphere equipped with a function that extends $R^{-1}\cos\varphi$. This function is smooth at the point at infinity $R = \infty$ of the sphere, and this point is noncritical for the extended function. Generic smoothing of the singularity at the origin transforms this extended function to a smooth function on the sphere with nondegenerate maximum and minimum points. In a neighborhood of the maximum point, the form d^2F is negative definite, and in a neighborhood of the minimum point, it is positive definite. Hence, generic smoothing creates at least two parabolic curves; one of them bounds the domain in which d^2F is positive definite and the other bounds the domain in which d^2F is negative definite.

Remark. We can readily construct an example of smoothing creating exactly two parabolic curves near the origin.

For smoothing of the form $F = f(R)\cos\varphi$ we have $f(R) = Rg(R^2)$. The curve of parabolic points is defined by the equation

$$\operatorname{tg}^2\varphi = -R^2\,\frac{f''}{f - Rf'} = 3 + 2R^2\,\frac{g''}{g'}, \quad \text{where } g' = \frac{dg}{d(R^2)},$$

as follows from formula (2). For $F = R(R^2 + \varepsilon)^{-1}\cos\varphi$, the parabolic curve defined by this equation has the form of a small lemniscate whose branches are intersecting transversally at the singular point. A small generic deformation merges the two hyperbolic quadrants and generates two parabolic ovals.

§5. Topological Properties of Gradient Mappings

Consider a smooth mapping A of an oriented plane onto an oriented plane. The domain in which the Jacobian J of the mapping A is positive is said to be *elliptic* and the domain where it is negative is said to be *hyperbolic*. The elliptic and hyperbolic domains of a generic mapping A are separated by a smooth parabolic curve $J = 0$.

Example. If $A = \operatorname{grad} F$, then the elliptic (resp., hyperbolic) domain is the domain over which the points of the graph of F are elliptic (resp., hyperbolic).

Definition. By the *index* $i(\gamma)$ of an oriented curve γ that does not intersect the parabolic curve of the mapping A we mean the number of revolutions of the tangent vector of the curve $A(\gamma)$ (for brevity we do not indicate the mapping A in this notation).

Example. The index of the boundary of a small disk belonging to the elliptic domain is $+1$, and that for the hyperbolic domain is -1. Let γ be a simple closed curve on the oriented plane oriented as the boundary of the disk bounded by the curve. Suppose that the curve does not intersect the parabolic curve.

Definition. By the *reduced index* of the curve we mean the number $\tilde{i}(\gamma) = i(\gamma) - \operatorname{sign}(J|\gamma)$.

Example. The reduced index of a small simple curve is zero.

Theorem 5. *The reduced index is additive. Namely, if M is a plane domain in which $J \neq 0$, and if M is bounded outside by a curve γ_0 and inside by curves γ_j ($j \geq 1$) (so that $\partial M = \gamma_0 - \sum \gamma_j$), then $\tilde{i}(\gamma_0) = \sum \tilde{i}(\gamma_j)$.*

Proof. A constant (translation-invariant) vector field v on the image plane can be lifted to the domain M immersed in the plane by the mapping A. The lift w has no zeros on M. Denote by ν_k and μ_k the numbers of revolutions of the tangent vector $\dot{\gamma}$ of the boundary and of the vector w along the component γ_k of the boundary of the domain M (orienting this component as the boundary of the disk bounded by this component). Then $\mu_0 = \sum \mu_j$ (since the field w has no zeros on M), while $\nu_k = 1$ for all k. Therefore,

$$\mu_0 - \nu_0 + 1 = \sum(\mu_j - \nu_j + 1). \tag{8}$$

The index of the curve γ is the number of revolutions of the vector $A_*(\dot{\gamma})$ with respect to the constant field v. This number differs from the number $\nu - \mu$ of the revolutions of the vector $\dot{\gamma}$ with respect to w only by the sign s, namely, $i(\gamma) = (\nu - \mu)s$. Hence, formula (8) proves that the reduced index $\tilde{i} = (\mu - \nu - 1)s$ is additive. □

Definition. A point at which the kernel of the derivative of the mapping A is tangent to the parabolic curve is called *special*.

For a generic mapping A, the order of tangency at a special point is one.

Definition. A special point is called *positive*, or *elliptic* (resp., *negative* or *hyperbolic*) if the vectors of the kernels pointing into the elliptic domain at neighboring parabolic points are directed from (resp., to) the parabolic point.

Example. For a gradient mapping $A = \operatorname{grad} F$, a special point is elliptic (resp., hyperbolic) depending on the nature of singularity of the field obtained from the field of asymptotic directions of the graph of F by lifting this field of asymptotic direction to the surface, in the fibered three-space of the tangent directions of the graph, defined by the field itself (regarded as the image of a section of the bundle of the tangent directions). This surface (representing the field of asymptotic directions) doubly covers the hyperbolic domain with a fold singularity of the projection over the parabolic curve (for a generic function F). The lifted direction field turns out to be (locally) the direction field of a smooth vector field that vanishes only over the special point. If the special point is elliptic, then the index of this vector field at its singular point is equal to $+1$ (i.e., this point is either a node or a focus), and if it is hyperbolic, then it is equal to -1 (a saddle point). Near an elliptic special point, the mapping A sends the hyperbolic half-neighborhood (in which the Jacobian is negative) to a small domain bounded by the cusp of the caustic (as a local diffeomorphism) and strongly stretches the elliptic half-neighborhood. At a hyperbolic point, the hyperbolic half-neighborhood is strongly stretched.

Consider two curves close to a parabolic curve γ in a plane, namely, the result γ^+ of a small shift of γ into the elliptic domain and the result γ^- of a small shift of γ into the hyperbolic domain (the three curves being oriented as the boundaries of the corresponding disks).

Theorem 6. *The difference between the indices of the curves γ^+ and γ^- is equal to the number of special points on the parabolic curves counted according to their signs:*

$$i(\gamma^-) - i(\gamma^+) = \sigma(\gamma)s(\gamma),$$

where $\sigma(\gamma)$ is the difference between the numbers of elliptic and hyperbolic special points, while the sign s is $+1$ if the parabolic curve is the exterior boundary of an elliptic domain (and -1 if it is the exterior boundary of a hyperbolic domain).

Corollary. *For the reduced indices we have $\tilde{i}(\gamma^-) - \tilde{i}(\gamma^+) = 2 + \sigma(\gamma)s(\gamma)$.*

Example. The only parabolic curve of the mapping $A(z) = \bar{z}^2 + 2z$ is the unit circle $|z| = 1$. The three special points are negative. In this case we have $i(\gamma^+) = 1$, $i(\gamma^-) = -2$, $\sigma = -3$, $s = +1$, $\tilde{i}(\gamma^+) = 0$, and $\tilde{i}(\gamma^-) = -1$.

Proof of the theorem. The curves $A(\gamma^-)$ and $A(\gamma^+)$ differ only by small spirals close to the special points (that is, near the cusps of the caustic, which is the curve of critical values of the mapping). Each spiral contributes ± 1 to the index. The spiral is placed on γ^+ if the special point is positive and on γ^- if it is negative. The contribution of the spiral to the index is equal to $-s$. This can readily be verified by treating, for instance, the four possible cases (and using the above description of the local structure of the mapping in neighborhoods of elliptic and hyperbolic special points; namely, the spiral is the image of the curve lying in the half-neighborhood of the special point that is strongly stretched). $\square$

Denote the boundary of a compact connected hyperbolic domain M of a gradient mapping by $\partial M = \gamma_0 - \sum \gamma_j$ (where γ_0 is the exterior boundary and the γ_j are interior boundaries, and all curves are oriented as the boundaries of disks). Denote by γ^+ the result of a small shift of the parabolic curve into the elliptic domain.

Theorem 7. *The reduced indices of the boundaries of the hyperbolic domain of a gradient mapping that are shifted into elliptic domains are additive, $\tilde{i}(\gamma_0^+) = \sum \tilde{i}(\gamma_j^+)$.*

Proof. The field of asymptotic directions (i.e., zeros of the quadratic form d^2F) is well defined on the double of the hyperbolic domain. The singular points of this field are provided by the special points of the boundary. The Euler characteristic of the double is $2-2n$, where n is the number of interior components, γ_j, of the boundary. Hence, the sum of indices of singular points of the field on the double is $\sigma(\gamma_0) + \sum \sigma(\gamma_j) = 2 - 2n$.

However, according to the corollary to Theorem 6 we have $\tilde{i}(\gamma_0^+) = \tilde{i}(\gamma_0^-) - 2 + \sigma(\gamma_0)$ because $s(\gamma_0) = -1$, and $\tilde{i}(\gamma_j^+) = \tilde{i}(\gamma_j^-) - 2 - \sigma(\gamma_j)$ because $s(\gamma_j) = 1$. We finally obtain

$$\tilde{i}(\gamma_0^+) - \sum \tilde{i}(\gamma_j^+) = \left[\tilde{i}(\gamma_0^-) - \sum \tilde{i}(\gamma_j^-)\right] + \left[2n - 2 + \sigma(\gamma_0) + \sum \sigma(\gamma_j)\right].$$

The first bracket vanishes by Theorem 5 and the second also vanishes, as was proved above. $\square$

Corollary. *The reduced indices of the parabolic curves of a gradient mapping that are shifted into the elliptic domain are zero.*

Proof. By the *rank* of a parabolic curve we mean the number of terms of the maximal (the longest) chain of parabolic curves successively embedded in one another and starting from the given curve (thus, the rank of a parabolic curve bounding an elliptic or hyperbolic disk is equal to 1). For curves of rank one bounding hyperbolic disks, the corollary readily follows from Theorem 7 (since the interior boundary is empty). For curves of rank one bounding elliptic disks, this follows from Theorem 5 (because the interior boundary is empty again). Theorem 7 represents the reduced index of a parabolic curve shifted into the elliptic domain as the sum of the reduced indices of parabolic curves of lower rank that are shifted into the elliptic domain. This proves the corollary by induction. $\square$

Remark. Panov and Aicardi, to whom I had communicated the above theory, observed that for any curve γ in the elliptic domain of a gradient mapping A the numbers of revolutions of the tangent vectors to the curves γ and $A(\gamma)$ coincide. This follows from the fact that the angle between the vectors $\dot{\gamma}$ and $A_*\dot{\gamma}$ never attains the values $\pm\pi/2$ (since the scalar product $(\dot{\gamma}, A_*\dot{\gamma})$ is of fixed sign in the elliptic domain).

Remark. Consider a gradient mapping of the plane and assume that this mapping is hyperbolic outside a sufficiently large disk. Theorem 7 shows that no information on the intrinsic structure of elliptic domains penetrates through the most exterior parabolic curves. The restrictions concerning the indices and the special points provided by the behavior of the mapping outside a large disk are related only to the indices of the most exterior parabolic curves and to the special points on these curves. From this point of view, the domains bounded by these exterior curves can be regarded as elliptic domains. I do not know whether one can really make these domains elliptic by deforming the function F.

Let us apply the above results to a generic gradient mapping of the plane that coincides with the nongeneric mapping $A(z) = \bar{z}^2$ (which has a D_4^- singularity at the origin) outside a large disk.

Theorem 8. *Every gradient mapping coinciding with A outside a disk centered at the origin has at least three special points on one of its most exterior parabolic curves.*

Proof. The index of a large circle $i(\gamma_0^-)$ is -2. Hence, its reduced index $\tilde{i}(\gamma_0^-)$ is -1. By Theorem 5, it is equal to the sum of the reduced indices of hyperbolic shifts of the exterior parabolic curves γ_j, and hence $\sum \tilde{i}(\gamma_j^-) = -1$. According to the corollary of Theorem 6, we have $\tilde{i}(\gamma_j^-) = \tilde{i}(\gamma_j^+) + 2 + \sigma(\gamma_j)$ $(s(\gamma_j) = 1$ because this parabolic curve is the exterior boundary of an elliptic domain). By the corollary of Theorem 7 we have $\tilde{i}(\gamma_j^+) = 0$. Thus, $\sum(2 + \sigma(\gamma_j)) = -1$. Hence, at least one of the summands is negative. For the sum of the signs of the special points on the exterior parabolic curve that corresponds to this summand we have the inequality $\sigma(\gamma_j) \leq -3$. Therefore, there are at least three special points on this parabolic curve. $\square$

Remark. Moreover, we have proved that there is an exterior parabolic curve on which the number of hyperbolic special points exceeds that of elliptic ones at least by three. This bound is sharp: for the gradient deformation $A(z) = \bar{z}^2 + \varepsilon z$ (provided by the versal deformation of the D_4^- singularity), there are exactly three special points (exactly three cusps on the caustic).

One may conjecture (in accordance with the general principle of topological minimality for algebraic objects) that similar inequalities hold for the other boundary conditions as well, namely, the set of singularities into which a degenerate gradient Lagrangian singularity is decomposed under a small generic gradient deformation that makes this singularity generic, is possibly topologically necessary (cannot be reduced by a different choice of a gradient deformation). For instance, in the two-dimensional case, consider the number of cusps of the caustic of a gradient mapping obtained by a continuation of a potential function (defining a degenerate singularity at the origin) from a small annulus centered at the origin to the interior disk. One may conjecture that the minimal value of this number of cusps is attained under a generic small deformation of the potential function, and that the other continuations, however sophisticated, provide a larger number of cusps.

Example. The degenerate singularity of the gradient mapping $z \mapsto \bar{z}^m$ has a standard small gradient deformation $z \mapsto \bar{z}^m + \varepsilon z$ with a single parabolic curve carrying $m + 1$ (hyperbolic) special points. *There exists no deformation* (*vanishing outside a disk centered at the origin*) *whose caustic has fewer cusps.*

Proof. As in Theorem 8 we see that $i(\gamma_0) = -m$, $\tilde{i}(\gamma_0^-) = 1 - m$, $\sum \tilde{i}(\gamma_j^-) = 1 - m$, and $\tilde{i}(\gamma_j^-) = \tilde{i}(\gamma_j^+) + 2 + \sigma(\gamma_j)$. If $-\sigma(\gamma_j) < m + 1$ for every exterior boundary, then $\sum \tilde{i}(\gamma_j^-) = \sum(2 + \sigma(\gamma_j)) > \sum(2 - m - 1) \geq 1 - m$, which contradicts the boundary condition $\sum \tilde{i}(\gamma_j^-) = 1 - m$. Thus, there exists an exterior parabolic curve for which $\sigma(\gamma_j) \leq -1 - m$. $\square$

Let us apply the above results to the study of the compactly supported deformations of a surface that is hyperbolic at infinity (over the Möbius band complementary to the disk on which the deformations are concentrated) and is transversal to the vertical direction. An example of such a surface at infinity is defined by the function

$$H(r,\varphi) = \cos 3\varphi + r\sin 2\varphi - cr^2\cos\varphi, \qquad 0 < c < 1/4, \tag{9}$$

in the polar coordinates of §2. We can readily verify the following properties of the field of asymptotic directions of a surface hyperbolic at infinity (we project the field to the Möbius band along the vertical direction and consider it as a field defined on the Möbius band). 1. *The two fields of asymptotic directions on the Möbius band are distinguishable*: one of them is orientable while the other is not. 2. *The number of the tangency points of asymptotic direction on the Möbius band with the central line at infinity is at least three*, where the number of the points of tangency with the orientable field is odd and with the nonorientable is even. This follows from the Möbius theorem, since the tangency points are exactly the projections of the inflection points of the line of intersection of the surface with the vertical projective plane containing the line at infinity. There are three such inflection points provided that the plane intersects the surface transversally along a smooth noncontractible curve (the perturbation can be so large that the perturbed surface can be tangent to some vertical lines). 3. Consider the torical surface fibered into circles over the double covering of the line at infinity, where the fibers consist of the unit vectors tangent to the cylinder that doubly covers, of the Möbius band at the central circle points (covering the points of the line at infinity). This bundle is induced by the restriction of the unit tangent vector bundle of the Möbius band by the mapping of the circle onto the line at infinity covering it twice. Each of the two fields of asymptotic directions along the line at infinity can be lifted to a section of the above bundle. *The intersection indices of the resulting curves on the torus with the section obtained by the lifting of the direction of the line at infinity are zero.* 4. To describe the rotation type of both asymptotic directions at infinity, it is convenient to use cyclic words of length $2(2k+1)$ ($2k+1$ being the total number of tangencies of both fields with the direction of the line at infinity). The four symbols of the alphabet describing the rotations are $\{\cup, \cap, \diagup, \diagdown\}$. Each symbol describes the rotation of a frame of asymptotic directions between two consecutive points of tangency of the asymptotic field with the central circle (on the cylinder, which doubly covers the Möbius band). The symbol "$\diagup$" ("$\diagdown$") stands for a positive (negative) rotation of the frame (the orientation of the cylinder being transported from the affine part of the projective plane). The grammar rules reflect the continuity of the rotation: "$\cup$" can be followed only by "$\cap$" and "$\diagup$", "$\cap$" only by "$\cup$" and "$\diagdown$", "$\diagup$" only by "$\cap$" and "$\diagup$", "$\diagdown$"only by "$\cup$" and "$\diagdown$". The cyclic words are supposed to be odd (the opposite symbols being opposite), and the number of the "$\cup$"s (and hence of the "$\cap$"s) should be odd. Since our words are cyclic, the words obtained from one another by the involution replacing each symbol by the opposite symbol can be identified. The symmetry allows us to represent our cyclical words by their halves, which I shall call *snakes*. Each snake is a word of $2k+1$ symbols, the total number of the "$\cup$"s and "$\cap$"s in a snake must be odd, and the snake produces a cyclic word of length $2(2k+1)$ if we follow it by a snake of opposite symbols. Opposite snakes represent the same cyclic word. A snake is said to be *admissible* if the corresponding cyclic word satisfies the above rules, and we do not distinguish the snakes representing the same cyclic word. For instance, there are only two admissible snakes of length 3, namely, $\cap\cup\cap$ and $\cap\setminus\setminus$. The numbers of admissible snakes of lengths $1, 3, 5, 7, 9, 11, \dots$ form a strange sequence $1, 2, 4, 10, 30, 94, \dots$. These numbers count the cyclically ordered partitions of an odd number n into an odd number of summands.

Remark. The numbers of unordered partitions of an odd number $n = 2k+1$ into an odd number of summands form a sequence $1, 2, 4, 8, 16, 29, \dots$ ($k = 0, 1, 2, 3, 4, 5, \dots$), which is even more strange. On classifying the partitions by the number of the summands, we obtain a triangle for which I could not find a general rule:

$$\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 4 & & 2 & & 1 & & \\
 & 1 & & 7 & & 5 & & 2 & & 1 & \\
1 & & 10 & & 10 & & 5 & & 2 & & 1
\end{array}$$

It seems that the admissible snakes are realizable by the fields of asymptotic directions of some surfaces that are hyperbolic at infinity, but I have not proved this. I do not know how a snake at infinity restricts the field of asymptotic directions at the parabolic curves over an affine disk for a surface which is hyperbolic at infinity. I think that computer experiments can be useful to obtain better understanding of this problem. Passing from the Möbius band (forming a neighborhood of the line at infinity) to the complementary affine disk, we obtain the following boundary conditions for the field of asymptotic directions ($d^2F = 0$) of the corresponding gradient mapping ($A = \operatorname{grad} F$). A) *The number of revolutions, of any of the two asymptotic directions, performed while the observation point makes one revolution along the boundary of a large disk on the affine plane is* $+1$ (provided that the surface is hyperbolic at infinity). In particular, *both fields of asymptotic directions are orientable over this large circle.* B) *The number* $i(\gamma_0^-)$ *of revolutions of the tangent vector to the image* $A\gamma_0^-$ *of a large circle, under the gradient mapping defined by a surface hyperbolic at infinity, is* $+1$. C) *The number of revolutions of the acceleration vector* $d^2A\gamma_0^-(\varphi)/(d\varphi)^2$ *performed while the observation point makes one revolution along the boundary of a large disk is at least four.* For our special example of the surface (9), this number of revolutions is four. In this example, the curve $A\gamma_0^-$ described by the image of the point moving along the circle $R = \text{const}$ bounding a large disk is a small deformation of a hypocycloid with three cusps passed twice in the negative direction. The deformation replaces three consecutive cusps by small spirals and smoothes the next three cusps. The tangent vector to the deformed curve makes $-2 + 3 = +1$ revolutions as the argument of the point running along the large circle changes from 0 to 2π. For cubic surfaces, the number of parabolic curves is equal to four and the number of special points is equal to six, and all these points are negative. This agrees with our above formulas,

$$\tilde{i}(\gamma_0^-) = \sum \tilde{i}(\gamma_j^-) = \sum (2 + \tilde{i}(\gamma_j^+) + \sigma(\gamma_j)) = 8 - 6 = 2 = i(\gamma_0^-) + 1.$$

Remark. It seems that a large positive value of the index $i(\gamma)$ (where γ is a curve in the hyperbolic domain bounding a disk containing the parabolic curves) implies the existence of many exterior parabolic curves of a given gradient mapping. This would follow if the Panov conjecture were true; this conjecture claims that, for a parabolic curve bounding a disk whose exterior points (close to the parabolic curve) are hyperbolic, the index $i(\gamma^-)$ of a small shift of the parabolic curve into the hyperbolic domain that is exterior for this curve is at most one. Panov stated this conjecture only for void parabolic curves bounding elliptic disks, but I know no counterexample even to the more general conjecture formulated above. This conjecture seems to be related to the following circumstances: 1) the index of a singular point of a plane gradient field is at most one; 2) the relation between the function and the direction of its gradient is invariant under the conformal diffeomorphisms; 3) every disk is conformally equivalent to the standard one; 4) all special points on a convex parabolic curve bounding an elliptic domain are negative. Of course, these arguments do not prove the conjecture, but the attempt to prove it by contracting the parabolic curve to a singular point of the surface seems to be natural.

References

1. B. Segre, The Nonsingular Cubic Surfaces, Clarendon Press, Oxford, 1942.
2. A. Hurwitz, "Über die Fourierschen Konstanten integrierbarer Funktionen," Math. Ann., **57**, 425–446 (1903).
3. V. I. Arnold, "Topological problems of the theory of wave propagation," Usp. Mat. Nauk, **51**, No. 1, 3–50 (1996).

Translated by V. I. Arnold