

On a new problem in Integral Geometry related to boundary problems for partial differential equations

Boris Paneah (Technion, Haifa)

1 Introduction

This paper is mainly devoted to a typical problem in integral geometry: to reconstruct a function in a domain D knowing its integrals over a family of subdomains in D . A peculiarity of the present work is that we consider bounded domains D with boundary ∂D . The statement of the problem under consideration and the results we obtain turn out to be intimately connected with both local and global properties of ∂D . Roughly speaking, given a connected bounded domain $D \subset \mathbb{R}^2$, let D_q be a system of subdomains in D , parameterized by points $q \in \partial D$. To any function f in \bar{D} we associate the integrals

$$\int_{D_q} f d\sigma = h(q), \quad q \in \partial D. \quad (1)$$

In connection with this equality, the above-mentioned problem in integral geometry can be formulated in the following way (see [1]): for which spaces of functions f and h is the map $f \mapsto h$ one-to-one, and which functions $h(q)$ can be represented by the integral (1)?

In this paper we investigate the situation in which D is a curvilinear plane triangle. In the framework of a class of function spaces we give an exhaustive answer to the above questions. With each of these function spaces we associate a noncommutative semigroup of maps in ∂D . This semigroup naturally generates a dynamical system with a set \mathcal{O} of orbits in ∂D . As we will show the solvability of equation (1) and the uniqueness of a solution depend only on whether the set \mathcal{O} contains cyclic orbits of a special type. Every such cycle plays the role of an obstruction when constructing a solution of equation (1).

The results we obtain turn out to be immediately applicable in Partial Differential Equations. In particular, we show in Sec. 7 that, generically, a unique solution of an arbitrary 3rd order strictly hyperbolic differential equation with constant coefficients in a bounded plane domain D can be reconstructed knowing its value on the *whole* boundary ∂D . And this is contrary to a common belief that all the boundary problems for hyperbolic equations with data on the *whole* boundary of a bounded domain are not well-posed.

In the course of obtaining the main results we transform equation (1) to an equivalent functional equation on the boundary ∂D . The solvability of such an equation turns out to be closely connected with the structure of a special semigroup of maps in ∂D . We establish the necessary and sufficient condition for the existence of a unique solution of the equation in question. In particular, the corresponding proof includes a new maximum principle for similar functional equations. We note that the functional equations under consideration have never been investigated before, and they are interesting by themselves.

Some preliminary results are contained in the author's papers [2,3].

2 Statement of the problem and definitions.

2.1 Let \mathbf{l}_1 and \mathbf{l}_2 be smooth nonsingular transversal vector fields in a disk $B \subset \mathbb{R}^2$. Introduce a curvilinear triangle $D = OA_1A_2$ whose sides OA_1 and OA_2 are trajectories of vector fields \mathbf{l}_1 and \mathbf{l}_2 , respectively. As to the side $\Gamma = A_1A_2$ it is assumed to be an arbitrary smooth curve without singularities which is transversal to OA_1 and OA_2 ¹. In addition the closure \bar{D} of a domain D is supposed to satisfy the following hypotheses.

1° For any point $p \in \bar{D}$ a trajectory of \mathbf{l}_j passing through p meets OA_k , $k, j = 1, 2$, $k \neq j$ at a point $\pi_j p$.

2° The set \bar{D} is \mathbf{l}_j -convex, $j = 1, 2$. This means that given points $p \in D$ and $q \in D$ on any trajectory γ_j of the field \mathbf{l}_j all the points $r \in \gamma_j$ between p and q belong to \bar{D} .

Given an arbitrary point $q \in \Gamma$ let D_q be a curvilinear parallelogram qq_1Oq_2 , where $q_j = \pi_j q$, $j = 1, 2$. The above conditions 1° and 2° guarantee an inclusion $\bar{D}_q \subset \bar{D}$ for all $q \in \Gamma$.

In this work we study a solvability of an integral equation of the following form:

$$\int_{D_q} f d\sigma = h(q), \quad q \in \Gamma. \quad (2)$$

Here $d\sigma$ is a measure in B , $h(q) \in C(\Gamma)$ is a given function and $f \in C(\bar{D})$ is an unknown function. The analogy of problem (2) with the famous Radon problem is obvious. It is immediately verified (with the help of equality (22) below, for instance) that the range of the operator

$$\mathcal{B} : f \in C(\bar{D}) \rightarrow \int_{D_q} f d\sigma \in C(\Gamma)$$

is (a part of) a linear space

$$\mathcal{H}(\Gamma) = (C^2 \cap C_0)(\Gamma)$$

of all twice continuously differentiable in Γ functions vanishing on the boundary of Γ . Therefore the best possible solution of the problem as it is formulated in Introduction consists in description of spaces $\mathcal{F}(D) \subset C(\bar{D})$ such that the map

$$\mathcal{B} : \mathcal{F}(D) \rightarrow \mathcal{H}(\Gamma)$$

is one-to-one. Among various candidates for the role of $\mathcal{F}(D)$ we have chosen a wide class of subspaces in $C(\bar{D})$ which naturally appear in the theory of boundary problems for PDE.

Definition Given a smooth nonsingular vector field \mathbf{l} in B we denote by $C_{\langle \mathbf{l} \rangle}(D)$ the set of all functions in $C(\bar{D})$ which remain constant along any trajectory of the field \mathbf{l} .

It is not difficult to describe the space $C_{\langle \mathbf{l} \rangle}(D)$ directly. Let

$$\mathbf{l} = \{\lambda_1(x), \lambda_2(x)\}, \quad x = (x_1, x_2) \in B$$

be a coordinate form of the vector field \mathbf{l} . Denote by $\omega(x)$ a smooth function without critical points which solves the first order differential equation

$$\lambda_1(x)(\partial/\partial x_1)\omega + \lambda_2(x)(\partial/\partial x_2)\omega = 0. \quad (3)$$

Then the space $C_{\langle \mathbf{l} \rangle}(D)$ consists of all functions

$$f(x) = F \circ \omega(x)$$

where F is an arbitrary continuous function on the range of ω .

2.2 Let \mathbf{l} be a smooth nonsingular vector field in B such that

¹See Fig. 1 on page 3.

- (i) \mathbf{l} is transversal to Γ and to both fields \mathbf{l}_1 and \mathbf{l}_2 ;
- (ii) for any point $p \in \overline{D}$ a trajectory of \mathbf{l} passing through p meets the curve Γ at a point $\pi_1 p$.

In virtue of hypothesis (ii) we are able to introduce the two maps in Γ

$$\zeta_1 = \pi_1 \circ \pi_1 \quad \text{and} \quad \zeta_2 = \pi_1 \circ \pi_2$$

which play a crucial role when formulating the main results of the work. These maps generate a noncommutative semigroup Φ_ζ . The elements of Φ_ζ are maps in Γ of the form

$$\zeta_J = \zeta_{j_n} \circ \dots \circ \zeta_{j_1}$$

where $J = (j_1, \dots, j_n)$ is an arbitrary multi-index with all j_k equal 1 or 2. The semigroup Φ_ζ naturally generates a dynamic system. In what follows we use the following geometric terminology.

- 1) Given a map $\zeta_J \in \Phi_\zeta$ an ordered set (q_1, \dots, q_{n+1}) of points in Γ is called a J -orbit if

$$q_{k+1} = \zeta_{j_k} q_k \quad \text{for} \quad 1 \leq k \leq n \leq \infty. \quad (4)$$

A set of all J -orbits will be denoted by \mathcal{O}_ζ .

- 2) If $q_1 = q_{n+1}$ or, equivalently, if $\zeta_J q_1 = q_1$, then the orbit (q_1, \dots, q_{n+1}) is called *cyclic* or in short, *cycle*.

Introduce the *critical sets*

$$\mathcal{T}_{\zeta_j} = \{q \in \Gamma \mid \mathbf{l}_j(q) \in T_q(\Gamma)\}, \quad j = 1, 2,$$

where $T_q(\Gamma)$ stands for the tangent space of the curve Γ at the point q . Let $\mathcal{T}_\zeta = \mathcal{T}_{\zeta_1} \cup \mathcal{T}_{\zeta_2}$.

- 3) If all the points of a J -orbit belong to the set \mathcal{T}_ζ then this orbit is called *critical*.
- 4) A J -orbit (q_1, \dots, q_{n+1}) is said to be \mathcal{T}_ζ -proper if in (4)

$$\zeta_{j_k} = \zeta_2 \quad \text{when} \quad q_k \in \mathcal{T}_{\zeta_1} \quad \text{and} \quad \zeta_{j_k} = \zeta_1 \quad \text{when} \quad q_k \in \mathcal{T}_{\zeta_2}.$$

From the point of view of the dynamical system generated by the semigroup Φ_ζ , in moving along any \mathcal{T}_ζ -proper orbit we leave each point $q_j \in \mathcal{T}_\zeta$ along a trajectory transversal to Γ .

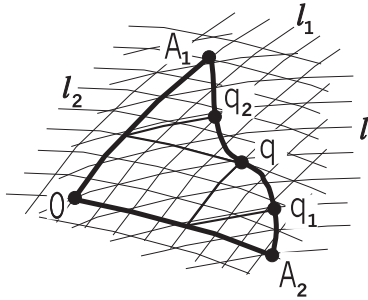


Figure 1

On this figure families of trajectories of the vector fields $\mathbf{l}_1, \mathbf{l}_2$ and \mathbf{l} are represented. The point $q \in \Gamma$ is in \mathcal{T}_{ζ_2} . Therefore the orbit (q, q_1) is proper whereas the orbit (q, q_2) is not.

Definition We denote by \mathfrak{N}_ζ the set of all critical \mathcal{T}_ζ -proper cycles in Γ .

3 The main results

We are now able to formulate the main results of this work. The first one treats the constant vector fields \mathbf{l} . Under small restrictions on the domain D this result solves the problem of the existence of the inverse operator

$$\mathcal{B}^{-1} : \mathcal{H}(\Gamma) \rightarrow C_{\langle \mathbf{l} \rangle}(D).$$

Theorem 1 Let \mathbf{l}_1 and \mathbf{l}_2 be constant vector fields in \mathbb{R}^2 . Let D be a domain as defined in Subsec.2.1 and at least one of the sets \mathcal{T}_{ζ_j} , $j = 1, 2$, is finite. Take an arbitrary constant vector field $\mathbf{l} = \lambda_1 \mathbf{l}_1 + \lambda_2 \mathbf{l}_2$ with positive λ_1 and λ_2 . Then given an arbitrary function $h \in \mathcal{H}(\Gamma)$ there exists a unique solution $f \in C_{(\mathbf{l})}(D)$ of equation (2) if and only if

$$\mathfrak{N}_{\zeta} = \emptyset. \quad (5)$$

The inverse operator $\mathcal{B}^{-1} : h \mapsto f$ is continuous: $\mathcal{H}(\Gamma) \rightarrow C_{(\mathbf{l})}(D)$.

To illustrate this result let us consider the following two figures.

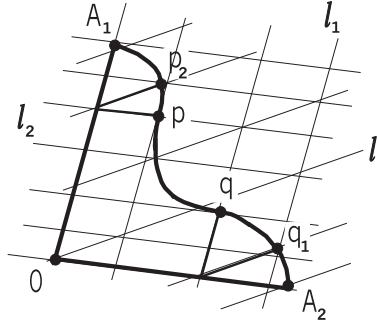


Figure 2

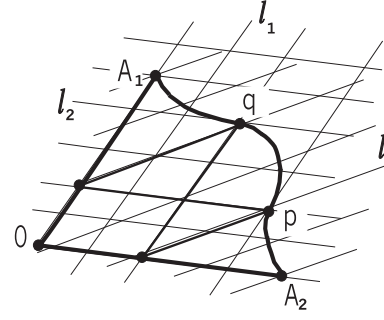


Figure 3

On both figures p is the only point in Γ belonging to \mathcal{T}_{ζ_1} and q is the only point in Γ belonging to \mathcal{T}_{ζ_2} . Therefore the only orbit consisting of the points p and q may belong to \mathfrak{N}_{ζ} . On Fig. 2 a unique \mathcal{T}_{ζ} -proper orbit beginning at the point p (q , respectively) contains as a second one the point $p_2 \notin \mathcal{T}_{\zeta}$ ($q_1 \notin \mathcal{T}_{\zeta}$, respectively). Thus $\mathfrak{N}_{\zeta} = \emptyset$, and problem (2) is uniquely solvable in the triangle OA_1A_2 for all functions $h \in \mathcal{H}(\Gamma)$.

On Fig.3 $\mathfrak{N}_{\zeta} \neq \emptyset$ as the J -orbit (p, q, p) with $J = (1, 2)$ is obviously in \mathfrak{N}_{ζ} . In view of Theorem 2 (see relation (42) below) the equation $\mathcal{B}f = h$ is not solvable if

$$\partial_s h(p) + \partial_s h(q) \neq 0.$$

Our subsequent results relate to nonconstant vector fields \mathbf{l} . At first we formulate rather general conditions on \mathbf{l} under which the hypothesis (5) guarantees the existence of the above operator \mathcal{B}^{-1} and its continuity. From the technical point of view this is the central and the most difficult part of the work.

Theorem 2 Assume that in addition to hypotheses in Theorem 1 related to $\mathbf{l}_1, \mathbf{l}_2$ and D a vector field $\mathbf{l} = (\lambda_1(x), \lambda_2(x))$ satisfies the following conditions:

1°

$$\partial_1(\lambda_1/\lambda_2) \geq 0, \quad \partial_2(\lambda_2/\lambda_1) \geq 0 \text{ in } D; \quad (6)$$

2° there is a solution ω of equation (3) such that for $j = 1, 2$

$$\partial_j \omega \partial_j^2 \omega \leq 0 \text{ in } D. \quad (7)$$

If $\mathfrak{N}_{\zeta} = \emptyset$, then equation (2) has a unique solution $f \in C_{(\mathbf{l})}(D)$ for an arbitrary function $h \in \mathcal{H}(\Gamma)$. The inverse operator \mathcal{B}^{-1} is continuous.

Although hypothesis (7) has an implicit nature the obtained result can be used in searching of various classes of desirable vector fields \mathbf{l} . In particular, the sufficiency of condition (5) in Theorem 1 (the most essential part of this Theorem) is a direct consequence of Theorem 2 as any constant vector field \mathbf{l} satisfies both hypotheses (6) and (7).

One more class of vector fields \mathbf{l} which guarantee the existence of a corresponding operator \mathcal{B}^{-1} can be easily found with the help of Theorem 2.

Theorem 3 Let $\mathbf{l}_1, \mathbf{l}_2$ and D be the same as in Theorem 1. Assume that a vector field \mathbf{l} satisfies the following condition: there are positive functions $\mu(x_2)$ and $\nu(x_1)$ such that

$$(i) \quad \lambda_1/\lambda_2 = \mu(x_2)/\nu(x_1); \quad (ii) \quad \partial_2\mu \leq 0, \quad \partial_1\nu \leq 0.$$

Then the conclusions of Theorem 2 remain valid provided that $\mathfrak{N}_\zeta = \mathfrak{o}$.

All these assertions will be proved in the subsequent sections.

We do not discuss here an application of the formulated results to boundary problems for partial differential equations (PDE) (see concluding Theorem 13). This requires a familiarity with a special terminology in PDE. This is why we postpone the corresponding discussion to Sec. 7 devoted exceptionally to boundary problems.

4 Proof of Theorem 2

The proof of Theorem 2 is long and consists of three parts. At first we reduce problem (1) *in equivalent manner* to two different functional-integral equations on the curve Γ . For the first one we obtain a maximum principle from which the uniqueness of a solution to problem (2) (i.e. the equality $\dim \ker \mathcal{B} = 0$) follows immediately. As to the second equation on Γ we prove that the spectral radius of the corresponding linear operator in $C(\Gamma)$ is less than one. This makes it possible to apply the F.Riesz-Schauder theory and to establish that the index of this operator equals zero. Therefore $\text{ind } \mathcal{B} = 0$. Combining this result with the uniqueness already proved completes the proof of Theorem 2.

4.1 Reduction of problem (2) to functional - integral equations

Without loss of generality we can restrict ourselves to unique vector fields \mathbf{l}_1 and \mathbf{l}_2 parallel to coordinate x_1 - and x_2 - axis, respectively. This makes it possible to consider the domain D as a curvilinear triangle whose sides are intervals $0 \leq x_1 \leq 1$ and $0 \leq x_2 \leq 1$ of x_1 - and x_2 - axis, respectively, and a curve

$$\Gamma = \{(x_1, x_2) \mid x_1 = \alpha_1(z), \quad x_2 = \alpha_2(z); \quad z \in I_Z\}.$$

Here

$$I_Z = \{z \mid -1 \leq z \leq 1\}$$

and $\alpha_1(z), \alpha_2(z)$ are smooth functions such that

$$|\alpha_1'(z)|^2 + |\alpha_2'(z)|^2 = 1, \quad z \in I_Z. \quad (8)$$

We note that

$$\alpha_1(-1) = 0, \quad \alpha_2(-1) = 1, \quad \alpha_1(1) = 1, \quad \alpha_2(1) = 0, \quad (9)$$

and

a point $(\alpha_1(z_0), \alpha_2(z_0)) \in \Gamma$ belongs to the critical set \mathcal{T}_{ζ_j} , $j = 1, 2$, if and only if

$$\alpha_j'(z_0) = 0.$$

Introduce a notation

$$\alpha(z) = (\alpha_1(z), \alpha_2(z)) \quad z \in I_Z.$$

It is convenient to treat α as a map

$$\alpha : I_Z \ni z \rightarrow \alpha(z) \in \Gamma$$

which is invertible due to (8). We observe also that in virtue of our assumptions about the domain D (see 1° and 2° in Subsec. 2.1) the inequalities

$$\alpha'_1(z) \geq 0 \quad \text{and} \quad \alpha'_2(z) \leq 0, \quad z \in I_Z \quad (10)$$

hold. (This fact is precisely what makes it possible to translate an invariant geometric description of the above domain D into an analytic language). We will prove these inequalities (which are not quite trivial) in Appendix.

Let $\omega = \omega(x)$ be an (arbitrary) fixed solution of equation (3) such that $\text{grad} \omega \neq 0$ in \overline{D} . As $\lambda_1 > 0$ and $\lambda_2 > 0$ it follows from (3) that $(\omega_{x_1} \omega_{x_2})(x) < 0$ in \overline{D} . Therefore one can assume (multiplying ω by -1 if it is necessary) that

$$\omega_{x_1}(x) > 0, \quad \omega_{x_2}(x) < 0 \quad \text{in} \quad \overline{D}. \quad (11)$$

Introduce notation

$$\gamma_- = \inf_D \omega, \quad \gamma_+ = \sup_D \omega,$$

and let $I_T = \{t \mid \gamma_- \leq t \leq \gamma_+\}$. Denote by ω_Γ the restriction of the function ω to Γ . We will prove now that *the map $\omega_\Gamma : \Gamma \rightarrow I_T$ is surjective and invertible*. To this end consider a map

$$\Omega = \omega \circ \alpha : I_Z \rightarrow I_T.$$

Differentiating function $\Omega(z) = \omega(\alpha(z))$ and using inequalities (8),(10) and (11) we arrive at inequality

$$\Omega'(z) > 0, \quad z \in I_Z. \quad (12)$$

This results in invertibility of Ω . But $\omega \circ \alpha = \omega_\Gamma \circ \alpha$ and by virtue of invertibility of α the same is true with respect to the map ω_Γ . To prove the surjectiveness of ω_Γ take a point $p_- \in \overline{D}$ such that $\omega(p_-) = \gamma_-$. Then a trajectory of the vector field \mathbf{l} passing through p_- meet Γ at a point q_- (see hypothesis (ii) in Sec. 2). Therefore $\omega_\Gamma(q_-) = \gamma_-$. In just the same way we determine a point $q_+ \in \Gamma$ such that $\omega_\Gamma(q_+) = \gamma_+$. Thus

$$\gamma_- = \min \omega_\Gamma \quad \text{and} \quad \gamma_+ = \max \omega_\Gamma,$$

and this completes the proof of the assertion.

Return now to definitions in Sec. 2 and note that

$$\pi_1(x) = x_1, \quad \pi_2(x) = x_2$$

for any point $x = (x_1, x_2)$ in Γ . By the definition of ω and ζ_j we have

$$\omega(x_1, 0) = \omega(\zeta_1(x)) \quad \text{and} \quad \omega(0, x_2) = \omega(\zeta_2(x)).$$

Denote

$$\omega_1(x) = \omega(x_1, 0), \quad \omega_2(x) = \omega(0, x_2).$$

Then

$$\omega_j(\alpha(z)) = \omega(\zeta_j \circ \alpha(z)) \quad j = 1, 2. \quad (13)$$

We now introduce smooth maps in I_T

$$\delta_j = \omega_j \circ \alpha \circ \Omega^{-1}, \quad j = 1, 2.$$

which play an important role in the following. Recalling that $\Omega = \omega \circ \alpha$ and using (13) we find that

$$\delta_j = \omega_\Gamma \circ \zeta_j \circ \omega_\Gamma^{-1}, \quad j = 1, 2 \quad (14)$$

and

$$\omega_j \circ \alpha = \delta_j \circ \Omega. \quad (15)$$

All the needed properties of the maps δ_j are collected in the following lemma.

Lemma 4 (i) Both maps δ_j in I_T are nondecreasing functions. In addition δ_1 maps I_T onto $[\gamma_0, \gamma_+]$ and δ_2 maps I_T onto $[\gamma_-, \gamma_0]$, where $\gamma_0 = \omega(0, 0)$.

(ii) For all values t , $\gamma_- < t < \gamma_+$, the inequalities

$$\delta_2(t) < t < \delta_1(t) \quad (16)$$

hold. Besides $\delta_1(\gamma_+) = \gamma_+$ and $\delta_2(\gamma_-) = \gamma_-$.

(iii) For any $t \in I_T$

$$\delta_1'(t) + \delta_2'(t) > 0 \quad (17)$$

Proof: (i) We note first of all that

$$(d/dz)\omega_j(\alpha(z)) = \omega_{x_j}(\pi_j(\alpha(z))) \circ \alpha_j'(z).$$

Combining (10) and (11) results in inequality $(d/dz)(\omega_j \circ \alpha)(z) \geq 0$. This together with (12) and (15) leads to the monotonicity of δ_j . Furthermore, in order to describe the images of δ_j it suffices to note that in virtue of (12)

$$\Omega^{-1}(\gamma_-) = -1, \quad \Omega^{-1}(\gamma_+) = 1,$$

and to verify directly that

$$\omega_1 \circ \alpha(1) = \gamma_+, \quad \omega_2 \circ \alpha(-1) = \gamma_-, \quad \text{and} \quad \omega_1 \circ \alpha(-1) = \omega_2 \circ \alpha(1) = \gamma_0$$

(ii) Inequalities

$$\omega_2 \circ \alpha(z) < \omega \circ \alpha(z) < \omega_1 \circ \alpha(z)$$

hold due to (11). Substituting $\Omega^{-1}(t)$ for z results in (16).

(iii) Differentiating (15) we find that

$$\omega_{x_j}(\pi_j \circ \alpha(z))\alpha_j'(z) = (\delta_j' \circ \Omega)(z)\Omega'(z), \quad (18)$$

and it remains to use inequalities (11), (12) and (8).

Analogously to Sec. 2 one can construct a semigroup Φ_δ of maps in I_T of a form $\delta_J = \delta_{j_1} \circ \dots \circ \delta_{j_n}$ with $J = (j_1, \dots, j_n)$ and all j_k equal to 1 or 2. Each map δ_J generates a set of J -orbits in I_T consisting of points $(t_1, t_2, \dots, t_{n+1})$ in I_T , $n \geq 1$, where

$$t_{k+1} = \delta_{j_k} t_k, \quad k \geq 1. \quad (19)$$

The union of such J -orbits over all multi-indices J we denote by \mathcal{O}_δ .

In virtue of (14) we find that for an arbitrary multi-index J

$$\delta_J = \omega_\Gamma \circ \zeta_J \circ \omega_\Gamma^{-1}. \quad (20)$$

Let $\mathcal{T}_{\delta_j} = \{t \in I_T \mid \delta_j'(t) = 0\}$, $j = 1, 2$, be a set of *critical* points of the function δ_j and let $\mathcal{T}_\delta = \mathcal{T}_{\delta_1} \cup \mathcal{T}_{\delta_2}$. A J -orbit $(t_1, t_2, \dots, t_{n+1})$ in I_T is called \mathcal{T}_δ -*proper* if in (20)

$$\delta_{j_k} = \delta_2 \quad \text{when} \quad t_k \in \mathcal{T}_{\delta_1} \quad \text{and} \quad \delta_{j_k} = \delta_1 \quad \text{when} \quad t_k \in \mathcal{T}_{\delta_2}.$$

As in Sec. 2 we introduce a set \mathfrak{N}_δ of all critical \mathcal{T}_δ -proper cyclic orbits in I_T . The following assertion allows to reformulate hypothesis (5) in Theorem 2 in a coordinate form.

Lemma 5 *The ω -image of any (critical \mathcal{T}_ζ -proper) J -orbit in \mathcal{O}_ζ is a (critical \mathcal{T}_δ -proper) J -orbit in \mathcal{O}_δ and conversely, any (critical \mathcal{T}_δ -proper) J -orbit in \mathcal{O}_δ is an ω -image of a (critical \mathcal{T}_ζ -proper) J -orbit in \mathcal{O}_ζ . In particular the sets \mathfrak{N}_ζ and \mathfrak{N}_δ can be empty only simultaneously.*

Proof: Let $\zeta_j(q) = \hat{q}$ for some j . In other words the sequence (q, \hat{q}) is a j -orbit in Γ . Then $\omega(\hat{q}) = \omega \circ \zeta_j(q) = \delta_j \circ \omega(q)$ due to (14), and consequently $(\omega(q), \omega(\hat{q}))$ is a j -orbit in I_T . If $\delta_j(t) = \hat{t}$, then $\omega_\Gamma^{-1}\hat{t} = \zeta_j(\omega_\Gamma^{-1}(t))$, so that (t, \hat{t}) is the ω -image of a j -orbit $(\omega_\Gamma^{-1}(t), \omega_\Gamma^{-1}(\hat{t}))$ in Γ . Furthermore, due to (18) for any point $x_0 = \alpha(z_0)$ we have

$$\alpha'_j(z_0) = 0 \iff \delta'_j(\omega(x_0)) = 0. \quad (21)$$

and hence the ω -image of an arbitrary point $x_0 \in \mathcal{T}_{\zeta_j}$ belongs to \mathcal{T}_{δ_j} and vice versa. Consequently

$$(x^{(1)}, \dots, x^{(n)}) \in \mathfrak{N}_\zeta \iff (\omega x^{(1)}, \dots, \omega x^{(n)}) \in \mathfrak{N}_\delta.$$

In particular

$$\mathfrak{N}_\zeta = \emptyset \iff \mathfrak{N}_\delta = \emptyset.$$

This completes the proof of the lemma.

Turn now to equation (2). Assuming $d\sigma = dx$ and denoting $h(\alpha(z))$ by $\tilde{h}(z)$ we rewrite it in a coordinate form:

$$(\mathcal{B}F)(z) := \int_0^{\alpha_1(z)} \int_0^{\alpha_2(z)} F \circ \omega(x_1, x_2) dx_2 dx_1 = \tilde{h}(z), \quad z \in I_Z. \quad (22)$$

In virtue of (9) $\mathcal{B}F \in \mathcal{H}(\Gamma)$ for all functions $F \in C(\overline{D})$. It follows that for any function $\tilde{h} \in \mathcal{H}(I_Z)$ equation (22) is equivalent to each of the two equations

$$(d/dz)\mathcal{B}F = (d/dz)\tilde{h}, \quad z \in I_Z, \quad (23)$$

and

$$(d^2/dz^2)\mathcal{B}F = (d^2/dz^2)\tilde{h}, \quad z \in I_Z. \quad (24)$$

In a detailed notation the first of them has a form

$$\alpha'_1(z) \int_0^{\alpha_2(z)} F(\omega(\alpha_1(z), x_2)) dx_2 + \alpha'_2(z) \int_0^{\alpha_1(z)} F(\omega(x_1, \alpha_2(z))) dx_1 = \tilde{h}'(z). \quad (25)$$

Introduce a function

$$G(t) = \int_0^t F(s) ds, \quad t \in I_T,$$

and substitute G' for F in (25). Integrating by parts reduces the equation obtained to a functional - integral equation

$$\begin{aligned} & (\alpha'_1/\omega_{x_2} \circ \alpha + \alpha'_2/\omega_{x_1} \circ \alpha)G(\omega \circ \alpha)(z) - (\alpha'_1/\omega_{x_2}(\alpha_1, 0))G(\omega_1 \circ \alpha)(z) \\ & - (\alpha'_2/\omega_{x_1}(0, \alpha_2))G(\omega_2 \circ \alpha)(z) - \mathcal{K}G(z) = \tilde{h}'(z), \quad z \in I_Z, \end{aligned} \quad (26)$$

with the additional condition $G(0) = 0$. Here \mathcal{K} is a nonpositive operator from $C(I_Z)$ to $C(I_T)$ that is

$$G \geq 0 \iff \mathcal{K}G \leq 0. \quad (27)$$

The latter follows from the explicit form of \mathcal{K}

$$\begin{aligned} \mathcal{K}G(z) &= \alpha'_1(z) \int_0^{\alpha_2(z)} (G \circ \omega)(\alpha_1(z), x_2) (1/\omega_{x_2}(\alpha_1(z), x_2))_{x_2} dx_2 \\ &+ \alpha'_2(z) \int_0^{\alpha_1(z)} (G \circ \omega)(x_1, \alpha_2(z)) (1/\omega_{x_1}(x_1, \alpha_2(z)))_{x_1} dx_1 \end{aligned}$$

if to take into account inequalities (10), (11) and (7). Substituting $\Omega^{-1}(t)$ for z in (26) we arrive immediately at the equation

$$\begin{aligned} G(t) - \mu_1(t)(G \circ \delta_1)(t) - \mu_2(t)(G \circ \delta_2)(t) \\ = \mathcal{K}_1 G(\Omega^{-1}(t)) + (\tilde{h}' \circ \Omega^{-1})(t), \quad t \in I_T. \end{aligned} \quad (28)$$

Here $\mathcal{K}_1 = (a'_1/\omega_{x_2} + a'_2/\omega_{x_1})^{-1}\mathcal{K}$ is nonnegative operator from $C(I_Z)$ to $C(I_T)$ and

$$\mu_j(t) = k_j(t)\alpha'_j(\Omega^{-1}(t)), \quad j = 1, 2, \quad (29)$$

where

$$\left. \begin{aligned} k_1(t) &= \frac{\omega_{x_1}\omega_{x_2}(\alpha \circ \Omega^{-1})(t)}{\omega_{x_2}(\alpha_1, 0)(\alpha'_1\omega_{x_1} \circ \alpha + \alpha'_2\omega_{x_2} \circ \alpha)(\Omega^{-1}(t))} > 0, \\ k_2(t) &= \frac{\omega_{x_1}\omega_{x_2}(\alpha \circ \Omega^{-1})(t)}{\omega_{x_1}(0, \alpha_2)(\alpha'_1\omega_{x_1} \circ \alpha + \alpha'_2\omega_{x_2} \circ \alpha)(\Omega^{-1}(t))} < 0. \end{aligned} \right\} \quad (30)$$

Remark. If \mathbf{l} is a constant vector field then ω can be chosen as a linear function. But then $\mathcal{K}_1 = 0$ and (28) becomes a pure functional equation. To the best of the author's knowledge, functional equations of such kind have not yet been investigated, except when δ_1 and δ_2 are linear functions.

From the definition, it follows that

$$t \in \mathcal{T}_{\delta_j} \iff \mu_j(t) = 0, \quad j = 1, 2. \quad (31)$$

4.2 The uniqueness of a solution of equation (2)

The required uniqueness is a direct consequence of the following assertion which is interesting all by itself.

Lemma 6 (Maximum principle for a functional equation) *If $\mathfrak{N}_\delta = \emptyset$ and a critical set \mathcal{T}_{δ_1} (\mathcal{T}_{δ_2} , respectively) is finite, then any solution G of the homogeneous equation (28) takes its maximum at the point γ_+ (γ_- , respectively).*

Proof: We assume for definiteness that the set \mathcal{T}_{δ_1} is finite. Take an arbitrary solution G of homogeneous equation (28) and introduce a function

$$G_{\mathcal{M}}(t) = G(t) - \mathcal{M}, \quad \text{where } \mathcal{M} = \max_{I_T} G.$$

Denote

$$\mathfrak{M} = \{t \in I_T \mid G(t) = \mathcal{M}\}.$$

As any constant solves the homogeneous equation (28) (what can be verified directly), the same is true with respect to the function $G_{\mathcal{M}}$. Let $t_1 \in \mathfrak{M}$ and $t_1 \neq \gamma_+$. Applying equality (28) with $h' = 0$ to function $G_{\mathcal{M}}$ and substituting t_1 for t we arrive at the equality

$$-\mu_1(t_1)G_{\mathcal{M}} \circ \delta_1(t_1) - \mu_2(t_1)G_{\mathcal{M}} \circ \delta_2(t_1) = \mathcal{K}_1 G_{\mathcal{M}}(\Omega^{-1}(t_1)). \quad (32)$$

Due to (8), (10), (29) and (30) the inequalities

$$\mu_1 \geq 0, \quad \mu_2 \geq 0$$

hold. As $G_{\mathcal{M}} \leq 0$ in I_T the right hand side in (32) is nonpositive whereas the left hand side is nonnegative. Consequently

$$\mu_1(t_1)G_{\mathcal{M}} \circ \delta_1(t_1) + \mu_2(t_1)G_{\mathcal{M}} \circ \delta_2(t_1) = 0.$$

In virtue of (31) if $t_1 \notin \mathcal{T}_\delta$, then

$$G_{\mathcal{M}} \circ \delta_1(t_1) = G_{\mathcal{M}} \circ \delta_2(t_1) = 0,$$

or, equivalently,

$$\delta_1(t_1) \in \mathfrak{M} \quad \text{and} \quad \delta_2(t_1) \in \mathfrak{M}.$$

But if $t_1 \in \mathcal{T}_{\delta_1}$ ($t_1 \in \mathcal{T}_{\delta_2}$, respectively), then in virtue of (17)

$$\delta_2(t_1) \in \mathfrak{M} \quad (\delta_1(t_1) \in \mathfrak{M}, \text{ respectively}).$$

As δ_1 and δ_2 are maps in I_T we can apply this argument to the points $t_2 = \delta_1(t_1)$ and $t_2 = \delta_2(t_1)$ and arrive at a new point $t_3 = \delta_{j_2}(t_2)$ also belonging to \mathfrak{M} . It is clear that continuing this procedure we obtain a set of all \mathcal{T}_δ -proper orbits $(t_1, t_2, \dots, t_n, \dots) \in \mathcal{O}_\delta$ completely lying in \mathfrak{M} . To prove Lemma 6 it remains to verify that *if $\mathfrak{N}_\delta = \emptyset$, then given any point $t_1 \in I_T$ there exists a \mathcal{T}_δ -proper orbit $\mathcal{O} = (t_1, t_2, \dots)$ which converges to γ_+ , i.e. $\lim_{n \rightarrow \infty} t_n = \gamma_+$.*

Remark If $\mathcal{T}_{\delta_1} = \emptyset$ i.e. the curve Γ has no points with a tangent line parallel to x_2 -axis, then one can complete the proof in several words. Indeed, in this situation for any point $t_1 \in I_T$ the orbit $(t_1, \delta_1(t_1), \dots, \delta_1^n(t_1), \dots)$ is \mathcal{T}_δ -proper. Furthermore, in virtue of inequality (16) the sequence $\delta_1^n(t_1)$ increases when $n \rightarrow \infty$. Denote $\lim_{n \rightarrow \infty} \delta_1^n(t_1) = \nu$. It is clear that $\delta_1(\nu) = \nu$. But this means that $\nu = \gamma_+$.

The general case is more difficult as any \mathcal{T}_δ -proper orbit $(t_1, t_2, \dots, t_n, \dots)$ containing points t_k from both sets \mathcal{T}_{δ_1} and \mathcal{T}_{δ_2} does not increase due to the same inequality (16).

We begin with several assertions.

Proposition 7 *If $t \neq \gamma_0$ and $(t_1, t), (t_2, t)$ are two \mathcal{T}_δ -proper orbits, then $t_1 = t_2$.*

Proof: Assume that $t_1 \neq t_2$ and

$$t = \delta_{j_1}(t_1), \quad t = \delta_{j_2}(t_2).$$

In virtue of Lemma 1(i) $j_1 = j_2$. Denote a common value of these indices by j . As $\delta_j' \geq 0$ it follows from the equality $\delta_j(t_1) = \delta_j(t_2)$ that $\delta_j'(t) = 0$ for all values t , $t_1 \leq t \leq t_2$. But then $t_1 \in \mathcal{T}_{\delta_j}$, whence $j_1 \neq j$ as the orbit (t_1, t) is \mathcal{T}_δ -proper.

Remark From a geometrical point of view the latter proposition asserts that two different \mathcal{T}_δ -proper orbits can not *enter* at the same point (although such orbits can leave the same point t if $t \notin \mathcal{T}_\delta$).

Proposition 8 *If a cyclic orbit S is a part of a \mathcal{T}_δ -proper orbit $\mathcal{O} = (t_1, t_2, \dots)$, then $t_1 \in S$.*

Proof: If $t_1 \notin S$ we let $t_q, q \geq 2$, be the first point in \mathcal{O} , belonging to S , so that

$$S = (t_q, \dots, t_{q+n}), \quad t_{q+n} = t_q.$$

But then $t_{q-1} \neq t_{q+n-1}$, and applying Proposition 7 leads to a contradiction.

Let

$$\lambda = \sup\{t \mid t \in \mathcal{T}_{\delta_1}\}.$$

Proposition 9 *If $\lambda < \gamma_+$, then there is an integer ν such that the inequality*

$$\delta_1^\nu(t) > \lambda \tag{33}$$

holds for all $t \in I_T$.

Proof: As the function $\delta_1(t)$ increases when $t > \lambda$ the inequality in question holds with an arbitrary $\nu \geq 1$ for all points $t > \lambda$. To find a required ν for $t < \lambda$ we let

$$\Lambda = \min_{t \leq \lambda} (\delta_1(t) - t).$$

Then $\delta_1(t) \geq t + \Lambda$ for $t \leq \lambda$ so that $\delta_1(\gamma_-) \geq \gamma_- + \Lambda$, $\Lambda > 0$. It follows that if

$$(\lambda - \gamma_-)\Lambda + 1 \geq \nu \geq (\lambda - \gamma_-)/\Lambda,$$

then

$$\delta_1^\nu(\gamma_-) \geq \gamma_- + \nu\Lambda \geq \lambda.$$

As $\delta_1' \geq 0$ we arrive at inequality (33) for all t , $\gamma_- \leq t \leq \lambda$. This completes the proof of Proposition 9.

The ending of the proof of Lemma 6. We say that a \mathcal{T}_δ -proper orbit (t_1, t_2, \dots) is δ_1 -oriented if for all $k \geq 1$

$$t_k \notin \mathcal{T}_{\delta_1} \Rightarrow t_{k+1} = \delta_1(t_k).$$

It is obvious that any point $t_1 \in I_T$ defines uniquely a δ_1 -oriented orbit. We first note that due to hypotheses the number λ does not equal to γ_+ and hence inequality (33) holds with a constant ν . If $t_1 > \lambda$, then the orbit $\mathcal{O} = (t_1, \delta_1(t_1), \dots, \delta_1^n(t_1), \dots)$ is δ_1 -oriented and, as was shown (see Remark), $\lim_{n \rightarrow \infty} \delta_1^n(t_1) = \gamma_+$. Let $t_1 \leq \lambda$. If a δ_1 -oriented orbit $\mathcal{O} = (t_1, \dots)$ does not lead to the point t_1 , then it does not contain any cyclic suborbit due to Proposition 8. In this case as \mathcal{T}_{δ_1} is a finite set, there exists a number m such that all the points t_{m+1}, t_{m+2}, \dots of the orbit \mathcal{O} lie outside of \mathcal{T}_{δ_1} . But then $t_{m+\nu} = \delta_1^\nu(t_m) > \lambda$, and as above $t_k \rightarrow \gamma_+$ when $k \rightarrow \infty$.

Consider a concluding situation: $t_m = t_1$ for a number m . Then the cyclic suborbit $S = (t_1, \dots, t_m)$ of \mathcal{O} can not be critical due to the hypothesis $\mathfrak{N}_\delta = \emptyset$. Consequently S contains a point t_q , $1 \leq q \leq m-1$, which is not from \mathcal{T}_δ . Then we introduce a new point $\hat{t}_{q+1} = \delta_2(t_q)$ (in contrast to the point $t_{q+1} = \delta_1(t_q) \in S$). Consider now a δ_1 -oriented orbit $\mathcal{O}_1 = (\hat{t}_{q+1}, \hat{t}_{q+2}, \dots)$. Being \mathcal{T}_δ -proper this orbit has no common points with S . Indeed if \hat{t}_p is the first such point and $\hat{t}_p = t_r$, $1 \leq r \leq m$, then in virtue of Proposition 7 $\hat{t}_{p-1} = t_{r-1}$. But the latter is not possible. Due to Proposition 8 the \mathcal{T}_δ -proper orbit $t_q, \hat{t}_{q+1}, \hat{t}_{q+2}, \dots$ does not contain cyclic suborbits. Hence, as above, all the points \hat{t}_{q+p+m} , $m \geq 1$, of the orbit \mathcal{O}_1 starting with some number p have the form

$$\hat{t}_{q+p+m} = \delta_1^m(\hat{t}_{q+p}).$$

This means that the orbit \mathcal{O}_1 converges to γ_+ , and so does the the sewing orbit

$$\mathcal{O} = (t_1, \dots, t_q, \hat{t}_{q+1}, \dots, \hat{t}_{q+p}, \hat{t}_{q+p+1}, \dots).$$

This completes the proof of the Maximum Principle.

To prove the uniqueness of a solution of equation (22) turn to the homogeneous equation (22) and prove that the solution F is zero. Indeed, by definition $F'(t) = G(t)$ with G a solution of the homogeneous equation (28). In view of the above maximum principle we have

$$\max_{[\gamma_-, \gamma_+]} G = G(\gamma_+) \quad (\text{or} \quad \max_{[\gamma_-, \gamma_+]} G = G(\gamma_-)).$$

As equation (28) is linear replacing G by $-G$ we find that

$$\min_{[\gamma_-, \gamma_+]} G = G(\gamma_+) \quad (\text{or} \quad \min_{[\gamma_-, \gamma_+]} G = G(\gamma_-), \text{ respectively}).$$

Thus $G \equiv \text{const}$ and hence $F \equiv 0$. This completes the proof of the uniqueness in Theorem 2.

4.3 The existence of a solution of equation (2)

As we know the equation in question is equivalent to both equation (23) and (24). Having prove the uniqueness of a solution of equation (23) in Subsec. 4.2 we established at the same time the uniqueness theorem for equation (24). Consequently the required existence follows if we will prove that the *linear operator* $(d^2/dz^2)\mathcal{B}$ in $C(I_Z)$ is Fredholm one and its index equals zero. It is worth mentioning that the operator d^2/dz^2 is an isomorphism between the spaces $\mathcal{H}(I_Z) = (C^2 \cap C_0)(I_Z)$ and $C(I_Z)$. To represent equation (24) in an expanded form we differentiate equation (25). Changing $\tilde{h}''(z)$ for $\tilde{H}(z)$ we arrive after some identical transformations at a functional-integral equation

$$\begin{aligned} & \left[\alpha_1'^2(z) \frac{\omega_{x_1}(\alpha(z))}{\omega_{x_2}(\alpha(z))} + \alpha_2'^2(z) \frac{\omega_{x_2}(\alpha(z))}{\omega_{x_1}(\alpha(z))} + 2\alpha_1'\alpha_2'(z) \right] F \circ \omega(\alpha(z)) \\ & - \alpha_1'^2(z) \frac{\omega_{x_1}}{\omega_{x_2}}(\alpha_1(z), 0) F \circ \omega(\alpha_1(z), 0) - \alpha_2'^2(z) \frac{\omega_{x_2}}{\omega_{x_1}}(0, \alpha_2(z)) F \circ \omega(0, \alpha_2(z)) \\ & = \alpha_1'^2(z) \int_0^{\alpha_2(z)} F \circ \omega(\alpha_1(z), x_2) \left(\frac{\omega_{x_1}}{\omega_{x_2}}(\alpha_1(z), x_2) \right)_{x_2} dx_2 \\ & + \alpha_2'^2(z) \int_0^{\alpha_1(z)} F \circ \omega(x_1, \alpha_2(z)) \left(\frac{\omega_{x_2}}{\omega_{x_1}}(x_1, \alpha_2(z)) \right)_{x_1} dx_1 + \tilde{H}(z). \end{aligned}$$

Substituting $\Omega^{-1}(t)$ for z yields

$$F(t) - \rho_1(t)F \circ \delta_1(t) - \rho_2(t)F \circ \delta_2(t) = \mathcal{N}F(t) + H(t), \quad t \in I_T, \quad (34)$$

where \mathcal{N} is an integral operator

$$\begin{aligned} \mathcal{N} : F \mapsto & \alpha_1'^2(t) \int_0^{\alpha_2(t)} F \circ \omega(\alpha_1(t), x_2) \left(\frac{\omega_{x_1}}{\omega_{x_2}}(\alpha_1(t), x_2) \right)_{x_2} dx_2 \\ & + \alpha_2'^2(t) \int_0^{\alpha_1(t)} F \circ \omega(x_1, \alpha_2(t)) \left(\frac{\omega_{x_2}}{\omega_{x_1}}(x_1, \alpha_2(t)) \right)_{x_1} dx_1 + \tilde{H}(t), \end{aligned}$$

$H(t)$ stands for the function $\tilde{H}(\Omega^{-1}(t))$ and

$$\begin{aligned} \rho_1(t) &= \frac{\alpha_1'^2(z)\lambda(\alpha_1(z), 0)}{\alpha_1'^2(z)\lambda \circ \alpha(z) + \alpha_2'^2(z)(1/\lambda) \circ \alpha(z) - 2\alpha_1'\alpha_2'(z)} \Big|_{z = \Omega^{-1}(t)}, \\ \rho_2(t) &= \frac{\alpha_2'^2(z)(1/\lambda)(0, \alpha_2(z))}{\alpha_1'^2(z)\lambda \circ \alpha(z) + \alpha_2'^2(z)(1/\lambda) \circ \alpha(z) - 2\alpha_1'\alpha_2'(z)} \Big|_{z = \Omega^{-1}(t)} \end{aligned}$$

with $\lambda(x) = (\lambda_2/\lambda_1)(x)$ (See (3)).

Remark If the vector field \mathbf{l} is proportional to a constant one, then $\mathcal{N} = 0$ and (34) becomes purely functional equation

$$F - \rho_1 F \circ \delta_1 - \rho_2 F \circ \delta_2 = H$$

with respect to an unknown function $F \in C(I_T)$.

The integral operator \mathcal{N} , as it follows from the classical Arzela criterion, is a compact operator in $C(I_T)$. As to the coefficients $\rho_1(t)$ and $\rho_2(t)$ what is important is the following properties:

- (i) $\rho_1(t)\rho_2(t) \geq 0$, $t \in I_T$;
- (ii) $0 < \rho_1(t) + \rho_2(t) \leq 1$ for all $t \in I_T$ and
 $\rho_1(t) + \rho_2(t) < 1$ for $t \notin \mathcal{T}_\delta$;
- (iii) $\rho_j(t) = 0 \iff t \in \mathcal{T}_{\delta_j}$, $j = 1, 2$.

The first inequality in (ii) follows from (8),(10) and (6), and the second one is based on (6). The implication (iii) is a consequence of (21).

Introduce a linear operator L in $C(I_T)$

$$L : F \mapsto \rho_1 F \circ \delta_1 + \rho_2 F \circ \delta_2.$$

In virtue of (ii) the norm of this operator $\|L\|$ does not exceed 1.² If $\|L\| < 1$ then the operator $E - L$ in the left hand side of (34) (with E the identical operator in $C(I_T)$) is invertible. In this case the required Fredholm property as well as the equality $\text{ind}(d^2/dz^2)\mathcal{B} = 0$ follows from the well known

Proposition 10 *If R is an invertible operator in a Banach space B and \mathcal{N} is a compact operator in B , then the operator $R - \mathcal{N}$ in B is Fredholm one and $\text{ind}(R - \mathcal{N}) = 0$.*

The following Proposition relating to the case $\|L\| \geq 1$ is also well known.

Proposition 11 *If L is a linear operator in B and $\|L^m\| < 1$ for some integer m , then the operator $E - L$ is invertible.*

For the completeness both Propositions are proved in Appendix.

From what has been said above it follows that the required existence of a solution to equation (34) (and consequently to equation (2)) is a direct consequence of the following assertion which can be considered therefore as a main result of this subsection.

Lemma 12 *If $\mathfrak{N}_\delta = \emptyset$ and at least one of the critical sets \mathcal{T}_{δ_1} and \mathcal{T}_{δ_2} is finite then there is an integer m such that $\|L^m\| < 1$.*

Proof: We note that for an arbitrary integer $N > 0$ the function $L^N f$ can be represented in a form

$$(L^N f)(t) = \sum_{j_1, \dots, j_N=1}^2 \rho_{j_1}(t) \rho_{j_2}(\delta_{j_1}(t)) \cdots \rho_{j_N}(\delta_{j_{N-1}} \circ \cdots \circ \delta_{j_1}(t)) f(\delta_J(t)), \quad (35)$$

where $J = (j_1, j_2, \dots, j_N)$. Indeed, for $N = 1$ this is true. To apply the induction assume this equality to be valid for some N and prove that it is valid for $N + 1$. But this is evident as by virtue of the definition of L and due to (35) we have

$$\begin{aligned} L^{N+1} f(t) &= L(L^N f) = \sum_{j_0=1}^2 \rho_{j_0}(t) (L^N f)(\delta_{j_0}(t)) \\ &= \sum_{j_0, j_1, \dots, j_N=1}^2 \rho_{j_0}(t) \rho_{j_1}(\delta_{j_0}(t)) \cdots \rho_{j_N}(\delta_{j_{N-1}} \circ \cdots \circ \delta_{j_0}(t)) f(\delta_J(t)) \end{aligned} \quad (36)$$

with $J = (j_0, j_1, \dots, j_N)$.

Applying the triangle inequality we find with the help of (35) that for an arbitrary function $f \in C(I_T)$ and at any point $t \in I_T$ the inequality

$$|L^N f(t)| \leq \sum_{j_1, \dots, j_N=1}^2 \rho_{j_1}(t) \rho_{j_2}(\delta_{j_1}(t)) \cdots \rho_{j_N}(\delta_{j_{N-1}} \circ \cdots \circ \delta_{j_1}(t)) \|f\| \quad (37)$$

holds. Let us prove that for an arbitrary point $t \in I_T$ one can find an integer N and a number $\gamma < 1$ such that for all functions f with $\|f\| = 1$ the inequality

$$|L^N f(t)| < \gamma \quad (38)$$

²We consider a standard norm $\|\cdot\|$ in $C(I_T)$: $\|F\| = \max_{t \in I_T} |F(t)|$.

is valid. If $t \notin \mathcal{T}_\delta$ then $\rho_{j_1}(t) + \rho_{j_2}(t) < 1$ due to (ii), and hence inequality (38) follows with $N = 1$. Let $t \in \mathcal{T}_\delta$. When proving the Maximum Principle we established that under hypotheses of the lemma there is a multi-index $\hat{J} = (j_1, \dots, j_{N-1})$, $N \geq 2$, such that the corresponding \hat{J} -orbit $t, \delta_{j_1}(t), \dots, \delta_{\hat{J}}(t)$ is \mathcal{T}_δ -proper and $\delta_{\hat{J}}(t) \notin \mathcal{T}_\delta$. Let N be a minimal integer with such a property. Then by the definition of a \mathcal{T}_δ -proper orbit we have

$$\rho_{j_1}'(t) = 0, \quad \rho_{j_2}'(\delta_{j_1}(t)) = 0, \quad \dots, \quad \rho_{j_{N-1}}'(\delta_{j_{N-2}} \circ \dots \circ \delta_{j_1}(t)) = 0,$$

with $j_k' + j_k = 3$ for all k , and hence due to (37)

$$|L^N f(t)| \leq \rho_{j_1}(t) \rho_{j_2}(\delta_{j_1}(t)) \cdots \rho_{j_{N-1}}(\delta_{j_{N-2}} \circ \dots \circ \delta_{j_1}(t)) \left(\rho_1(\delta_{\hat{J}}(t)) + \rho_2(\delta_{\hat{J}}(t)) \right).$$

In virtue of (ii) the first $N - 1$ factors on the right hand side are not greater than 1 whereas the latter one is strictly less than 1. This completes the proof of inequality (38).

By continuity this inequality remains valid with the same N for all points in a neighborhood of the point t . Choosing a finite covering of I_T by such neighborhoods we will find a number m for which inequality (37) is valid with $N = m$ for all points $t \in I_T$. In doing so it should be taken into account that, according to (36), if inequality (38) is true with some N it remains valid for all numbers $N' > N$. This completes the proof of Lemma 12 and hence the proof of the existence in Theorem 2.

To complete the proof of Theorem 2 it remains to show that the inverse operator \mathcal{B}^{-1} (whose existence follows from the uniqueness of a solution to equation (2)) is bounded: $\mathcal{H}(\Gamma) \rightarrow C_{\langle \mathbf{1} \rangle}(D)$. But we have proved that this operator is defined on the whole Banach space $\mathcal{H}(\Gamma)$ (the existence part in Theorem 2). In order to use the Banach's closed graph theorem³ and thus to establish the desirable boundedness we have only to show that the operator \mathcal{B} is bounded: $C_{\langle \mathbf{1} \rangle}(D) \rightarrow \mathcal{H}(\Gamma)$. The latter is equivalent to an a priori estimate

$$\|\mathcal{B}F, \mathcal{H}(\Gamma)\| \leq c \|F, C(\overline{D})\|, \quad F \in C_{\langle \mathbf{1} \rangle}(\overline{D}),$$

with $c > 0$ a constant. But

$$\|\mathcal{B}F, \mathcal{H}(\Gamma)\| = \max_{I_Z} |(d^2/dz^2)\mathcal{B}F| + \max_{I_Z} |\mathcal{B}F|,$$

and making use of the explicit form of the operator $(d^2/dz^2)\mathcal{B}$ (see the equality preceding (34)) we arrive easily at the latter inequality. This completes the proof of Theorem 2.

5 Proof of Theorem 1

As any constant vector field \mathbf{l} satisfies conditions (6) and (7), using Theorem 2 we conclude that if $\mathfrak{N}_\zeta = \emptyset$ then equation (2) has a unique solution $f \in C_{\langle \mathbf{1} \rangle}(D)$ for an arbitrary function $h \in \mathcal{H}(\Gamma)$. In just the same way a continuity of the inverse operator \mathcal{B}^{-1} follows from Theorem 2.

To prove the necessity of hypothesis (5) assume that $\mathfrak{N}_\zeta \neq \emptyset$. Let $(q_1, q_2, \dots, q_{n+1})$ be a critical \mathcal{T}_ζ -proper cyclic J -orbit in Γ with $J = (j_1, j_2, \dots, j_n)$. According to Lemma 5 there is a critical \mathcal{T}_δ -proper cyclic J -orbit $(t_1, t_2, \dots, t_{n+1})$ in I_T responding to the same multi-index J . The latter means according to definitions (see Subsec. 4.1) that

$$\left. \begin{array}{l} \text{(i)} \quad t_1 = t_{n+1} \text{ (cyclicity)} \\ \text{(ii)} \quad t_{k+1} = \delta_{j_k}(t_k), \quad 1 \leq k \leq n, \\ \text{(iii)} \quad \delta_{j_k}'(t_k) = 0, \quad 1 \leq k \leq n, \end{array} \right\} \quad (39)$$

³This theorem states: if B_1 and B_2 are Banach spaces and $L : B_1 \rightarrow B_2$ is a continuous one-to-one linear operator on B_1 onto B_2 , then its inverse L^{-1} is also continuous.

where $j'_k = 3 - j_k$ is (in a sense) a complement index to j_k . Turn to equation (28), with H replacing $\tilde{h}' \circ \Omega^{-1}$. As \mathbf{I} is a constant vector field the corresponding function $\omega(x_1, x_2)$ can be chosen as linear one. But then both functions ω_{x_1} and ω_{x_2} are constants. Combining this observation with equalities (29) and (30) results in the identity

$$\mu_1(t) + \mu_2(t) = 1, \quad t \in I_T. \quad (40)$$

On the other hand, using the same reasoning the operator \mathcal{K}_1 equals 0. Thus equality (28) becomes

$$G(t) - \mu_1(t)G(\delta_1(t)) - \mu_2(t)G(\delta_2(t)) = \widehat{h}(t), \quad (41)$$

where each function $\mu_j(t)$, $j = 1, 2$, is proportional to $\delta'_j(t)$ what follows from (18). Let us substitute t for t_1 in (41). Using subsequently relations (39) and (40) we find that one of the numbers $\delta'_j(t_1)$, equals zero which means that $\mu_j(t_1) = 0$, whereas $\mu_{j'}(t_1) = 1$ and $\delta_{j'}(t_1) = t_2$. This results in the equality

$$G(t_1) - G(t_2) = \widehat{h}(t_1).$$

Using this procedure subsequently for $t = t_2, \dots, t = t_n$ and applying the cyclicity of the orbit in question on the last step we arrive at the system of equalities

$$\begin{aligned} G(t_1) - G(t_2) &= \widehat{h}(t_1), \\ G(t_2) - G(t_3) &= \widehat{h}(t_2), \\ &\dots \\ G(t_n) - G(t_1) &= \widehat{h}(t_n). \end{aligned}$$

Summing up these equalities we get

$$\sum_{j=1}^n \widehat{h}(t_j) = 0.$$

By the definition of \widehat{h} this equality is equivalent to the equality

$$\sum_{j=1}^n \partial_s h(q_j) = 0. \quad (42)$$

Here ∂_s is the differentiation with respect to the natural parameter on Γ . Thus each critical \mathcal{T}_ζ -proper cycle $(\zeta_1, \dots, \zeta_{n+1})$ generates a relation (42) involving any given function h in (2). A violation of this relation leads to the unsolvability of equation (2) with this h . In other words each orbit from \mathfrak{N}_ζ represents an obstruction when constructing the inverse operator \mathcal{B}^{-1} . Thus we have proved the part "only if" in Theorem 1. This completes the proof of this theorem.

Remark Note that the problem of the completeness of a system of the above obstructions as well as the problem whether or not these conditions are sufficient for the solvability of problem (2) remain still open.

6 Proof of Theorem 3

To prove this theorem it suffices to note that under conditions of the theorem the function

$$\omega = \int_0^{x_1} \nu(s) ds - \int_0^{x_1} \mu(t) dt$$

solves equation (3) and satisfies both conditions (6) and (7). Thus the result follows from Theorem 2.

7 First boundary problem for hyperbolic differential equations

As an application of the obtained results a new boundary problem for a wide class of hyperbolic differential operators in the plane will be studied in this section. The main distinctive features of this problem is that it is considered in a *bounded* domain, and the value of an unknown function is given on the *whole* boundary of the domain. In this connection it is worth mentioning that in the framework of the classical theory of PDE boundary problems for hyperbolic equations are usually considered in domains which are intimately connected with the corresponding equation (half space, half cylinder, an angle between characteristics in \mathbb{R}^2 etc). If the domain is bounded, a part of the boundary is usually free of a priori information about unknown solution. The evolutionary character of hyperbolic equations seems to impose a taboo on a priori information about a solution on the whole boundary of a bounded domain. However as we show below (see Theorem 13) for a wide class of hyperbolic equations this taboo can be lifted. In domains closely connected with the corresponding hyperbolic differential operators solutions of equations in question are uniquely defined if their values on boundaries of these domains are known (the first boundary problem).

7.1 Statement of the problem

For the sake of brevity we restrict ourselves to a homogeneous differential operator with constant coefficients.

In the (x, y) -plane \mathbb{R}^2 we consider an arbitrary homogeneous x -strictly hyperbolic operator $P(\partial_x, \partial_y)$ of the 3rd order. The x -strictly hyperbolicity means that the characteristic polynomial $P(\tau, \lambda)$ has, for any $\lambda \neq 0$, three distinct real roots in τ . It follows that the operator $P = P(\partial_x, \partial_y)$ can be uniquely represented in the form

$$P(\partial_x, \partial_y) = a(\partial_x - a_1\partial_y)(\partial_x - a_2\partial_y)(\partial_x - a_3\partial_y) \quad (43)$$

with some constants a, a_1, a_2, a_3 , where $a_j \neq a_k$ for $j \neq k$. The characteristics of the operator P are straight lines

$$y + a_1x = \text{const}, \quad y + a_2x = \text{const}, \quad y + a_3x = \text{const}.$$

Let $\mathbf{l}_1, \mathbf{l}_2$ and \mathbf{l}_3 be vector fields in \mathbb{R}^2 parallel to these lines, respectively. Denote by $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_6$ characteristic rays beginning at some point 0. Choose any triple of neighboring rays \mathcal{R}_j , say, $\mathcal{R}_1, \mathcal{R}_2$ and \mathcal{R}_3 . Let \mathcal{R}_3 be the ray lying between \mathcal{R}_1 and \mathcal{R}_2 . Consider a curvilinear triangle $D = OA_1A_2$ with sides $OA_1 \subset \mathcal{R}_1$, $OA_2 \subset \mathcal{R}_2$. As to the side $\Gamma = A_1A_2$ it is assumed to be an arbitrary smooth curve without singularities which is transversal to OA_1 and OA_2 (cf. Subsec 2.1). We suppose the closure \bar{D} to satisfy the hypotheses 1° and 2° of Subsec. 2.1. It follows in particular that Γ is transversal to the vector field \mathbf{l}_3 .

The *First Boundary Problem* for the above operator $P(\partial_x, \partial_y)$ and domain D is as follows.

Given functions $F \in C(\bar{D})$ and $h \in C(\partial D)$ find a solution of the boundary problem

$$Pu = F \quad \text{in } D, \quad u = h \quad \text{on } \partial D. \quad (44)$$

We call a function u in \bar{D} a *generalized solution* of the problem (44) if $u \in C^2(D)$, $u = h$ on ∂D , and for all functions $\varphi \in C_0^\infty(D)$

$$\int_{\mathbb{R}^2} u {}^tP\varphi dx dy = \int_{\mathbb{R}^2} F\varphi dx dy,$$

where tP is the formally adjoint differential operator.

7.2 The formulation of the result and a sketch of the proof

To formulate the main result concerning a solvability of problem (44) let us consider the semi-group Φ_ζ of maps in Γ introduced in Subsec. 2.2 with $\mathbf{l} = \mathbf{l}_3$. The critical sets \mathcal{T}_{ζ_j} considered in Subsec. 2.2 are now nothing but the sets of characteristic points in Γ with respect to the operator P and we call them *characteristic* sets. In exactly the same way we introduce J -orbits as well as the notions cyclic, critical and \mathcal{T}_{ζ_j} -proper orbits. Finally we introduce the set \mathfrak{N}_ζ whose elements are all the \mathcal{T}_{ζ_j} -proper cyclic orbits, consisting of only characteristic points in Γ .

Denote by $C^k(\partial D)$ the space of continuous on ∂D functions whose restrictions to all sides of the triangle D are k times continuously differentiable functions. The main result concerning the problem (44) is as follows.

Theorem 13 *Assume that at least one of the characteristic sets \mathcal{T}_{ζ_1} and \mathcal{T}_{ζ_2} is finite. Then for any functions $F \in C(\overline{D})$ and $h \in C^2(\partial D)$ there exists a unique generalized solution $u(x, y)$ of the problem (44) if and only if the set \mathfrak{N}_ζ is empty. The inverse operator $(F, h) \mapsto u$ is continuous: $C(\overline{D}) \times C^2(\partial \Omega) \rightarrow C^2(\overline{D})$. If $F \in C^k(D)$ and $h \in C^{k+1}(\partial D)$, $k \geq 1$ is an integer, then $u \in C^{k+2}(D)$ is a classical solution of the problem in question.*

Proof: We restrict ourselves to the proof of the existence of a unique generalized solution to problem (44) with $F = 0$. Let us write down the operator P in the form (43). It is obvious that there exists a linear transformation in \mathbb{R}^2 reducing the problem under consideration to the problem

$$(m_1 \partial_x + m_2 \partial_y) \partial_x \partial_y u = 0 \quad \text{in } D, \quad u = h \quad \text{on } \partial D. \quad (45)$$

Here D is a domain in \mathbb{R}^2 whose boundary ∂D consists of three parts $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, where

$$\Gamma_1 = \{(x, y) \mid y = 0, \quad 0 \leq x \leq 1\}, \quad \Gamma_2 = \{(x, y) \mid x = 0, \quad 0 \leq y \leq 1\},$$

$$\Gamma_3 = \{(x, y) \mid x = \alpha_1(t), y = \alpha_2(t); \quad 0 \leq t \leq 1\},$$

and

$$\alpha_1(0) = 0, \quad \alpha_1(1) = 1; \quad \alpha_2(0) = 1, \quad \alpha_2(1) = 0.$$

(For convenience we preserve the previous notations for the domain and functions).

Let

$$h = h_1(x) \quad \text{on } \Gamma_1, \quad h = h_2(y) \quad \text{on } \Gamma_2, \quad \text{and} \quad h = h_3(x, y) \quad \text{on } \Gamma_3.$$

The continuity of the function h leads to the natural compatibility conditions

$$h_1(0) = h_2(0), \quad h_1(1) = h_3(1, 0), \quad h_2(1) = h_3(0, 1). \quad (46)$$

Due to assumptions about the domain D an arbitrary generalized solution u to the equation in (45), satisfying boundary condition only on $\Gamma_1 \cup \Gamma_2$ can be represented in the form

$$u(x, y) = \int_0^x \left(\int_0^y F(n_1 s + n_2 t) dt \right) ds + h_1(x) + h_2(y) - h_1(0), \quad 0 \leq x, y \leq 1. \quad (47)$$

Here $n = (n_1, n_2)$ is a unit vector which is orthogonal to the vector $m = (m_1, m_2)$ and $n_1 > 0$, $n_2 < 0$. As to the function F , this is an arbitrary continuous function on the interval $I = (n_2, n_1)$. The necessity of satisfying the boundary condition $u = h_3$ on Γ_3 leads to the following integral equation for an unknown function $F \in C(I)$:

$$\int_0^{\alpha_1(t)} \left(\int_0^{\alpha_2(t)} F(n_1 x + n_2 y) dy \right) dx = H(t), \quad 0 \leq t \leq 1. \quad (48)$$

Here

$$H(t) = -h_1(\alpha_1(t)) - h_2(\alpha_2(t)) + h_3(\alpha_1(t), \alpha_2(t)) + h_1(0)$$

is a given function. What is important is that the function $H(t)$, generated by an arbitrary continuous and twice piecewise differentiable function h in (45), belongs to the space $\mathcal{H}(I) = C^2 \cap C_0(I)$ (see Subsec. 2.1). This follows from the compatibility conditions (46). Conversely, the function $u(x, y)$ which is defined by (47) with F a solution of equation (48) solves the problem (45).

Thus the problem (45) turns out to be equivalent to the equation (48) which is nothing but the equation (22). The existence of a unique solution to the problem (45) provided that $\mathfrak{N}_\zeta = \emptyset$ follows immediately from Theorem 1.

Appendix

Proposition A.1 *With the notation introduced in Subsec. 4.1 the inequalities*

$$\alpha'_1(z) \geq 0, \quad \alpha'_2(z) \leq 0, \quad z \in I_Z \tag{A.1}$$

hold.

Proof: It suffices to prove the first inequality.

(I) We note that if the curve Γ in a neighborhood of a point $M(x_0, y_0) \in \Gamma$ is described by an equation $y = f(x)$ with f a differentiable function, then one of half-intervals $\{(x, y) \mid x = x_0, 0 \leq y - y_0 \leq \varepsilon\}$ and $\{(x, y) \mid x = x_0, -\varepsilon \leq y - y_0 \leq 0\}$ is free of points of \overline{D} .

(II) We note also that given differentiable functions $\varphi_1(t)$ and $\varphi_2(t)$ with the same range there are points t_1 and t_2 such that

$$\varphi_1(t_1) = \varphi_2(t_2) \quad \text{and} \quad \varphi'_1(t_1)\varphi'_2(t_2) \neq 0. \tag{A.2}$$

Indeed, localizing the problem and using an affine transformation one can reduce the proof to the case when domains of $\varphi_1(t)$ and $\varphi_2(t)$ coincide and $\varphi'_1(t) > 0$. But then the result is obvious: as t_2 we take any point with $\varphi'_2(t_2) \neq 0$, and we choose $t_1 = \varphi_1^{-1} \circ \varphi_2(t_2)$.

(III) To prove the first inequality in (A.1) assume that $\alpha'_1(\tau) < 0$ for some $\tau \in I_Z$ and take a point $\hat{\tau}$ with $\alpha'_1(\hat{\tau}) > 0$. (As $\alpha_1(-1) = 0$ and $\alpha_1(1) = 1$ such a point certainly exists). We choose a point $\theta \in (\tau, \hat{\tau})$ for which

$$A := \alpha_1(\theta) = \max_{\tau \leq t \leq \hat{\tau}} \alpha_1(t).$$

It is clear that there are positive numbers θ_1, θ_2 such that

$$B := \alpha_1(\theta - \theta_1) = \alpha_1(\theta + \theta_2).$$

But then the restrictions $\alpha_-(t)$ and $\alpha_+(t)$ of the function $\alpha_1(t)$ to the intervals $(\theta - \theta_1, \theta)$ and $(\theta, \theta + \theta_2)$, respectively, map their domains on (A, B) . In view of (II) there exist points t^* and t_* such that

$$x^* = \alpha_-(t^*) = \alpha_+(t_*) \quad \text{and} \quad \alpha'_-(t^*)\alpha'_+(t_*) \neq 0. \tag{A.3}$$

We now consider three points in ∂D

$$M_0 = (x^*, 0), \quad M^* = (x^*, \alpha_2(t^*)), \quad M_* = (x^*, \alpha_2(t_*)).$$

For definiteness let $\alpha_2(t_*) > \alpha_2(t^*)$. Due to (A.3) each point M^* and M_* has a neighborhood in which Γ is described by an equation

$$x_2 = f(x_1), \quad |x - x^*| < r,$$

with $f = \alpha_2 \circ \alpha_1^{-1}$. In virtue of (I) on the straight line $x_1 = x^*$ a half-neighborhood of each point M^* and M_* is free of points of D . As M_0 and M^* belong to \overline{D} and \overline{D} is l_2 -convex, the line segment M_0M^* lies wholly in \overline{D} . But this contradicts to what has been said about the point M_* . This completes the proof of Proposition A.1 .

Proof of Proposition 10: It is clear that

$$\dim \ker(R - \mathcal{N}) = \dim(E - R^{-1}\mathcal{N}) < \infty \quad (\text{A.4})$$

since $R - \mathcal{N} = R(E - R^{-1}\mathcal{N})$. Denote $R - \mathcal{N} = P$ and $E - R^{-1}\mathcal{N} = Q$ so that $P = RQ$. By definition

$$\text{coker } Q = B/\mathcal{R}(Q) = \{f + \mathcal{R}(Q)\},^4$$

where $f + \mathcal{R}(Q)$ denotes a class of elements f' in B such that $f - f' \in \mathcal{R}(Q)$. Therefore

$$\text{coker } P = RB/\mathcal{R}(AQ) = \{Rf + \mathcal{R}(AQ)\}.$$

But $RB = B$ due to invertibility of R and $Rf - Rf' \in \mathcal{R}(AQ)$ if and only if $f - f' \in \mathcal{R}(Q)$. This means that

$$\dim \text{coker } P = \dim \text{coker } Q.$$

It remains to combine the latter equality with (A.4) and to use the compactness of the operator $R^{-1}\mathcal{N}$ and the F.Riesz-Schauder theorem⁵.

Proof of Proposition 11: It suffices to prove the convergence of the series $\sum_{k=0}^{\infty} \|R^k\|$. Let $\|T\| = \tau$ and $\|T^m\| = \gamma < 1$. Then for all integers $p = 1, \dots, m-1$ the inequality $\|T^{n+m+p}\| \leq \tau^p \gamma^n$ holds. This implies

$$\sum_{k=0}^{\infty} \|R^k\| = \sum_{n=0}^{\infty} \sum_{p=0}^{m-1} \|T^{n+m+p}\| \leq (\tau^m - 1)/(\tau - 1)(1 - \gamma).$$

References

- [1] I.Gel'fand, M.Graev and N.Vilenkin, *Integral Geometry and Representation Theory*, New York, Academic Press, 1966.
- [2] B.Paneah, On a Problem in Integral Geometry Connected to the Dirichlet Problem for Hyperbolic Equations, IMRN International Mathematics Research Notices, 1997, No.5.
- [3] B.Paneah, A Problem in Integral Geometry with Applications to the Dirichlet Problem for a Class of Hyperbolic Differential Equations, *Differential Equations, Asymptotic Analysis and Mathematical Physics*, Math.Res. 100, Akademie-Verlag, Berlin, 1997, 256-261.

⁴ $\mathcal{R}(Q)$ denotes the range of the operator Q .

⁵This theorem states: if E and \mathcal{N} are identical and compact operators (respectively) in a Banach space B , then $\text{ind}(E - \mathcal{N}) = \dim \ker(E - \mathcal{N}) - \dim \text{coker}(E - \mathcal{N}) = 0$.