Jahresber Dtsch Math-Ver (2011) 113:183–184
DOI 10.1365/s13291-011-0031-x

Vorwort Heft 4-2011

Hans-Christoph Grunau

Online publiziert: 12. Oktober 2011

Am 3. Juni 2010, neun Tage vor seinem 73. Geburtstag, verstarb in Paris Vladimir Igorevich Arnold. Begraben wurde er in Moskau, der Stadt, in der er die meiste Zeit gelebt und gearbeitet hat; seit 1993 war er allerdings zugleich auch an der Dauphine in Paris tätig. Lassen sich – nicht zuletzt bedingt durch Reisebeschränkungen seitens der sowjetischen Behörden – die Aufenthaltsorte Arnolds relativ genau eingrenzen, so gelingt dieses hinsichtlich seines Schaffens und seiner Wirksamkeit nun überhaupt nicht. Zwar konzentrieren sich Leonid Polterovich und Inna Scherbak – langjährige Teilnehmer an Arnolds Seminaren – im vorliegenden Nachruf auf Hamiltonsche Dynamik, symplektische Topologie und Singularitätentheorie, beigetragen hat er aber in mehr als 20 Büchern (die teilweise in mehreren Auflagen erschienen und in verschiedene Sprachen übersetzt wurden) und mehr als 300 Arbeiten zu zahlreichen Gebieten der Mathematik von der Algebraischen Geometrie über Differentialgleichungen, globale Analysis, Hydromechanik und statistische Mechanik hin zur Zahlentheorie, um nur einige zu nennen. Den größten Anteil an Arnolds Bekanntheit haben vielleicht seine Durchbrüche in der KAM-Theorie, einem grundlegenden Resultat aus der Theorie Hamiltonscher Systeme zur Existenz quasiperiodischer Lösungen, das nach Kolmogorov, ihm selbst und Moser benannt ist. Im Nachruf findet sich auch eine lange Liste der Arnold zuerkannten Preise, eine Nominierung für die Fieldsmedaille 1974 wurde offenbar nach Intervention der Sowjetunion nicht weiter verfolgt.

Die Buchbesprechungen diskutieren Neuerscheinungen zu der zeitabhängigen von Kármánschen Differentialgleichung für dünne elastische Platten sowie aus dem Bereich der mathematischen Statistik und hier insbesondere zur Fehleranalyse statisti-

H.-Ch. Grunau (✉)
Institut für Analysis und Numerik, Fakultät für Mathematik, Otto-von-Guericke-Universität, Postfach 4120, 39016 Magdeburg, Deutschland
e-mail: hans-christoph.grunau@ovgu.de

scher Schätzverfahren. Schließlich wird ein Buch zum Wechselspiel zwischen zellulären Automaten, Gruppen- und Ringtheorie vorgestellt, grundlegende Konzepte gehen hier auf das Werk John von Neumanns zurück.

Jahresber Dtsch Math-Ver (2011) 113:185–219
DOI 10.1365/s13291-011-0027-6

HISTORISCHER BEITRAG

V.I. Arnold (1937–2010)

Leonid Polterovich · Inna Scherbak

Received: 10 June 2011 / Published online: 7 September 2011

This photo is a shot from an amateur movie made by Prof. V. Lin (Technion, Haifa, Israel) at his home when he hosted Vladimir Arnold, on January 28, 1997

L. Polterovich · I. Scherbak (✉)
School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel
e-mail: scherbak@post.tau.ac.il

L. Polterovich
e-mail: polterov@post.tau.ac.il

L. Polterovich
Department of Mathematics, University of Chicago, 5734 S. University Ave, Chicago, IL 60637, USA
e-mail: polterov@math.uchicago.edu

Abstract This article is devoted to V.I. Arnold, a famous mathematician who passed away in June 2010. We discuss life and times of Arnold, and review some of his seminal contributions to symplectic geometry and singularities theory which were among Arnold's favorite subjects.

Keywords Symplectic manifold · Lagrangian submanifold · Hamiltonian system · Isolated singularity · Reflection group

Mathematics Subject Classification (2000) 01A70 · 57R17 · 70H08 · 14B05 · 11F22

1 Life and Times of V.I. Arnold

Those who know the material will not learn anything new, and those who do not know it will not understand anything.

V.I. Arnold about a badly written introduction

Vladimir Igorevich Arnold was born on June 12, 1937, in Odessa, USSR (now Ukraine), where his mother's family was living at that time. The family name, Arnold, apparently, comes from an 18th Century Prussian army officer who fled to Russia after killing his friend in a duel.

Arnold considered himself as a fourth-generation mathematician. His paternal great-grandfather was S.B. Zhitkov, who wrote a textbook on mathematics although he worked in a bank. His grandfather, V.F. Arnold, was a mathematical economist. His father, I.V. Arnold, was first a student and later a professor at the Moscow Lomonosov University. One of his students was a famous Russian nuclear physicist, dissident and human rights activist Andrei D. Sakharov who highly appreciated him as a teacher. I.V. Arnold passed away when Vladimir was eleven.

Vladimir I. Arnold grew up in Moscow in an intellectual atmosphere. Among his family's close friends and relatives where mathematicians, physicists, chemists and biologists. He recalled [20] "Probably the main result of my frequent childhood conversations with various outstanding scholars was a sense of the deep unity of all the sciences..."

Early on in his life, Vladimir took part in the activities of the Children's Scientific Society organized by "the father of Soviet cybernetics", A.A. Lyapunov, for the children of his friends. The weekly meetings of the Society, held at Lyapunov's place, were devoted to diverse subjects. The first talk the young Arnold ever presented was on wave interference. He was ten years old at the time. Arnold recollected [20] that among the young members of the society there were future members of the Soviet Academy of Sciences and a prominent cardiologist.

The participants also were fond of cross-country skiing. It became Arnold's favorite winter sport which he enjoyed his entire life. "When I get stuck with a problem I put on my cross-country skis, and when I get back after 30 or 50 kilometers of skiing, I have a new idea. If the idea does not work, I use the same method again".

Arnold first experienced "a joy of discovery" in 1949 when his "first genuine math teacher", an elementary school teacher I.V. Morozkin, posed a problem about two elderly ladies walking towards each other.[1] The 12 years old boy thought about this problem all the day, and then "suddenly the solution came as a revelation". "It was the desire to experience this joy of discovery again and again that made me a mathematician", Arnold told.

In 1954 he was admitted to the Moscow Lomonosov State University, despite the fact that his mother was Jewish, and that his maternal grandfather A.S. Isakovich was arrested in 1938. Arnold explained his admission by a combination of two factors: Stalin's death and the personal intervention of I.G. Petrovsky, a prominent mathematician, who at that time became a rector of the University.

Those were the "golden years" of the Mechmat (Department of Mechanics and Mathematics at Moscow University). Famous professors such as P.S. Aleksandrov, I.M. Gelfand, L.A. Lusternik, A.Ya. Khinchin, A.N. Kolmogorov, I.G. Petrovsky, L.S. Pontryagin were teaching exceptional students such as V.M. Alexeev, D.V. Anosov, V.I. Arnold, A.A. Kirillov, Yu.I. Manin, S.P. Novikov, Ya.G. Sinai.

Arnold became a student of A.N. Kolmogorov whom he admired all his life. As the first project, Kolmogorov suggested Arnold to work on Hilbert's thirteenth problem. Vladimir solved it in 1957 by using earlier results of Kolmogorov. He showed that any continuous function of several variables can be represented as a composition of a finite number of functions of two variables. This work became the subject of his Ph.D. (*Candidate of Sciences*) thesis which was defended in 1961. Two years before that he was awarded his M.Sc. degree for a thesis "On mappings of a circle to itself". In 1963 Arnold became a *Doctor of Sciences* (an analogue of Habilitation) for his seminal work on small denominators and stability problems in classical and celestial mechanics.

In 1961, Arnold became a Mechmat faculty member, and in 1965 he was promoted to Full Professor. In 1986 he moved to the Steklov Institute of Mathematics in Moscow (still keeping ties to the Mechmat until 1994). He worked there until his death in June 2010. From 1993 to 2005 he also held the position of Professor at Université Paris-Dauphine, spending Spring semesters in Paris and Falls in Moscow.

Arnold spent 1965 in Paris as a postdoctorate at the Sorbonne. At the request of his supervisor, J. Leray, Arnold delivered a one-semester course on dynamical systems. The audience included many renowned mathematicians (Cartan, Douady, Fréchet, Godement, Leray, Schwarz, Serre, Thom). One of the participants, Andre Avez, recorded the lectures and then published them as a book [22]. That year, Arnold also attended Thom's seminar on singularities at the IHES. He later admitted, "The meeting with R. Thom has greatly changed my mathematical world". "I learned singularity theory during my four-hour long conversation with B. Morin, after his remarkable talk on Whitney and Morin singularities at the Thom Seminar". Arnold considered R. Thom one of his mentors in Mathematics (along with Kolmogorov and Gelfand).

[1] Two elderly ladies left their home towns at sunrise heading towards each other. Each of them was walking at a constant speed. After meeting at noon they kept walking in the same directions, and reached the other ladies' home town at 4 p.m. and at 9 p.m. respectively. When did the sun come up that morning?

After returning to Moscow, Arnold started his own seminar mostly devoted to singularity theory. Every Tuesday, from October to December and from March to May, for more than 40 years, at 4 pm participants gathered in room 14–14 on the 14th floor of the main building of Moscow University. Participants included both undergraduate and graduate students as well as mature mathematicians. Interestingly enough, many of them had no formal connection to the mathematical establishment. Let us clarify this point. Moscow mathematical life[2] had the following structure: The official layer included the Moscow State University and the Steklov Institute, both with strong anti-Semitic leanings and strictly controlled by the Communist Party and the KGB. Numerous scientists with "Jewish roots" were doing mathematics as a hobby, in addition to their full time jobs as engineers and researchers in obscure industrial research institutes. Fortunately, there was also an unofficial layer, a kind of mathematical oasis, where these "outsiders" enjoyed the luxury of being supervised by several world-acclaimed gurus (Arnold, Gelfand, Manin, Novikov, Sinai). Arnold made the effort to turn his seminar into a great show. Seemingly speakers were the props, while Arnold was the star. But usually the speakers benefited from this arrangement because Arnold explained their own results to them, so that they could finally understand what they have proved. Each semester, at the first meeting of the Seminar, Arnold discussed and commented on some open problems.[3] Arnold's seminar became an *alma mater* for several generations of his students (A. Khovansky, A. Givental, A. Varchenko, V. Vassiliev among many others) which nowadays form *Arnold's School*.

Arnold had been denied permission to travel abroad since 1968, after he (along with 99 Soviet mathematicians) had signed a letter addressed to the Soviet authorities, protesting the psychiatric confinement of a notable dissident and mathematician Esenin-Volpin.[4] Arnold did not leave the Soviet Union until the beginning of the *perestroika* in the late 1980s.

Since 1993 some sessions of Arnold's Seminar were held in Paris, at the Jussieu Mathematical Institute, in addition to the usual Moscow Seminars. When Arnold was out of Russia, the Seminar's sessions would be organized by his former students S. Gussein-Zade, V. Zakalukin, and A. Khovansky. Sometimes Arnold sent letters to the Moscow branch containing problems and comments. In Moscow, Arnold's Seminar met until December 2010.

The breadth of Arnold's mathematical interests is breathtaking. He was among the founders of modern Hamiltonian dynamics (Kolmogorov–Arnold–Moser theory, Liouville–Arnold theorem, Arnold's diffusion), symplectic topology (Arnold's conjectures) and singularity theory. This will be discussed in more details in the next sections.

Arnold wrote a very influential paper [9] on real algebraic geometry, where he found new restrictions on ovals of plane algebraic curves by using methods of four-dimensional topology. He considered real algebraic geometry as a very important

---

[2] To be precise, we refer here to the 1980s.

[3] In 2000, the problems supplied with up-to-date comments from Arnold and some participants of the Seminar, were collected in the book "Arnold's Problems" Phasis, Moscow. (English translation: Springer-Verlag & Phasis, 2005).

[4] As a result of this letter, the story became public, and after the Voice of America gave a broadcast on the topic, Esenin-Volpin was soon released.

subject which is still far from being understood: “Unfortunately, algebraic geometers are unable to solve the real problems”.

He discovered [5, 26] a profound link between hydrodynamics and geometry of (infinite-dimensional) diffeomorphism groups. Being undoubtedly aware of the major significance of his discovery, Arnold joked in [12]: “The formulas … for the curvatures [of diffeomorphism groups] can be used even for rough estimates of the time over which a long-term dynamical prediction of the weather is impossible, if we agree to a few simplifying assumptions…: The earth has the shape of the torus.”

Arnold pioneered the study of cohomology of braid groups [8] and wrote an important paper “Modes and quasimodes” [10] on quasi-classical approximation.

He had a very distinctive style both as a researcher and a writer. His trademark was to explain main ideas and to show simple examples relating the issue to real world problems. He considered mathematics as an experimental science [19]: “Mathematics is a part of physics. Physics is an experimental science, a part of natural science. Mathematics is the part of physics where experiments are cheap.” Arnold used to run various mathematical experiments and, from time to time, he proudly showed to his students thick notebooks with calculations written in artistic handwriting. He hated formal axiomatically based exposition and always looked for a more intuitive geometric explanations. His texts often contained beautiful illuminating pictures. Some of them, such as Arnold’s famous cat illustrating a mixing automorphism of the two-dimensional torus became a logo of the mathematical subject. At the same time (in contrast to physicists!) Arnold mastered a coordinate-free language which enabled him to highlight the role of abstract mathematical structures involved.

Arnold was a mathematical time-traveler: in his studies he often went back to the classics, discovering there new insights and using them as a basis for new profound problems for future generations. For instance, Arnold’s famous conjectures on fixed points of symplectic diffeomorphisms appeared as a far-reaching generalization of Poincaré’s “last geometric theorem”. Arnold had a vivid interest in the history of mathematical ideas which manifested itself in his masterpiece “Huygens and Barrow, Newton and Hooke” [17].

He was one of the organizers of the Moscow Center for Continuous Mathematical Education. Since 2001, this center, in cooperation with the mathematical section of the Russian Academy of Sciences, has been organizing an annual summer school on contemporary mathematics in Dubna, a small town near Moscow, for about a hundred high school and undergraduate students. Arnold lectured there almost every year.

Arnold was very concerned about the state of mathematical education in primary and high schools. He criticized “the de-geometrization of mathematical education and the divorce from physics” taking place in Western countries, and did everything he could to prevent disastrous reforms of mathematical education in Russia.

According to Arnold, his first encounter with the axiomatic method was when he was eleven years old and tried to understand the multiplication of negative numbers [20]. Ever since then, he hated the axiomatic approach to teaching math based on “unmotivated definitions.” “It is only possible to understand the commutativity of multiplication by counting and re-counting soldiers by ranks and files or by calculating the area of a rectangle in the two ways. Any attempt to do without this interference by physics and reality in mathematics is sectarianism and isolationism which will destroy the image of mathematics as a useful human activity in the eyes of all sensible people." It is ridiculous to teach "addition of fractions to children who have never cut (at least mentally) a cake or an apple into equal parts. No wonder the children will prefer to add a numerator to a numerator and a denominator to a denominator". "The main goal of the mathematical education should be to cultivate the ability to mathematically investigate the phenomena of the real world" [19].

In 2004, Arnold wrote a booklet, "Problems for children from 5 to 15 years old", where he collected 79 problems intended "to develop a culture of thinking".

Arnold received various prestigious awards including the Lenin Prize (1965, with Kolmogorov) which was the highest award in the USSR, the Crafoord Prize (1982, with L. Nirenberg),[5] the Harvey Prize (1994), the Heineman Prize for Mathematical Physics (2001), the Wolf Prize in Mathematics (2001), the State Prize of the Russian Federation (2007) and the Shaw Prize in Mathematical Sciences (2008). However, he did not have the chance to compete for the Fields medal even though he was nominated for it in 1974: the Soviet authorities succeeded in their efforts to have his name withdrawn from the list of nominees.

Arnold became a member of the national academies of the USA and France, as well as of the Royal Society (UK). Later on, shortly before the collapse of the Soviet Union, he was at last elected to the Soviet Academy of Sciences.

In Spring 1999, while riding a bicycle (yet another favorite sport of his) in a forest near Paris, Arnold had an accident which resulted in a traumatic brain injury. In spite of the doctors' fears, he made a good recovery and returned to his mathematical and non-mathematical activity. Arnold died on June 3rd, 2010, in Paris, nine days before his 73rd birthday. He was buried in Moscow, in the Novodevichy cemetery.

Because of the enormous scope of Arnold's mathematical heritage, we could not deem to discuss all of his major achievements. Rather we focus on some of Arnold's contributions to symplectic topology and to singularity theory. The selection reflects our personal tastes. We have attempted to restrict references to textbooks and surveys where possible, with the exception for Arnold's original works.

In the past year a number of excellent memorial articles on Vladimir I. Arnold have appeared (see e.g. [38]). To some extent we are influenced by them and there are inevitable overlaps. We wave any claim to originality.

For reflections on various facets of Arnold's personality we refer the reader to the personal accounts [37, 48, 54]. But perhaps the most authentic source is Arnold's book [20] containing various recollections and stories which he used to tell.

In Sect. 2 we review some of Arnold's contributions to symplectic topology and Hamiltonian dynamics. In Sect. 3 we take a route of singularity theory: it leads, through algebraic and topological invariants of singularities, to a beautiful link between singularities, reflection groups and root systems discovered by Arnold in the early 1970's. We complete the paper with a discussion on interrelations between symplectic geometry and singularities.

[5] The Soviet authorities did not allow him to travel to Stockholm to receive it.

2 Symplectic Topology and Hamiltonian Dynamics

By symplectic topology I mean the discipline having the same relation to ordinary topology as the theory of Hamiltonian dynamical systems has to the general theory of dynamical systems.

V.I. Arnold, First steps in symplectic topology

2.1 Mathematical Model of Classical Mechanics

Let us start with a brief description of the mathematical model of classical mechanics. The reader is invited to consult Arnold's textbook [12] for further details. This ground-breaking book (which appeared in Russian in 1974) was among the very first ones worldwide which systematically presented classical mechanics in the language of symplectic geometry.

Consider the motion of a mass m particle in the configuration space $\mathbb{R}^n$ (equipped with the coordinate q) in the field of a potential force $F = -\frac{\partial U}{\partial q}$. The dynamics of the particle is governed by Newton's 2nd law $m\ddot{q} = F$. (Here and below $\dot{q}$ stands for the velocity of the particle and $\ddot{q}$ for its acceleration.) This second order ODE turns out to be quite complicated and with rare exceptions it cannot be solved explicitly. Therefore a qualitative theory is needed. First of all let us make a little trick and introduce the auxiliary momentum variable $p = m\dot{q}$. Let $H(p,q)$ be the full energy of the particle,

$$H(p,q) = \frac{1}{2}m|p|^2 + U(q),$$

where the first term on the right-hand side stands for the kinetic energy and the second one for the potential energy. With this notation Newton's second law can be rewritten as a *Hamiltonian system* of first order ODE's

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} = -\frac{\partial H}{\partial q} \end{cases} \tag{1}$$

in the *phase space* $\mathbb{R}^{2n}$ equipped with the coordinates p and q. The evolution of the particle is given by a family of diffeomorphisms $h_t : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ which send an initial condition $(p(0), q(0))$ to the solution $(p(t), q(t))$ at time t.

This reformulation paves a way for a qualitative analysis of the particle motion which in fact is applicable to any energy function $H(p,q,t)$, in general time-depending and with a "nice" behavior at infinity. We refer to H as to a *Hamiltonian function*, and we call $\{h_t\}$ the *Hamiltonian flow* generated by H. When H does not depend on time, the system (1) readily yields the energy conservation law: the Hamiltonian H is constant along the trajectories.

A basic feature of Hamiltonian flows is that they preserve the volume form $\sigma = dp_1 \wedge dq_1 \wedge \cdots \wedge dp_n \wedge dq_n$ in the phase space. This statement, known as the Liouville theorem, gave rise to a powerful mathematical abstraction of a mechanical system, namely to the notion of an automorphism of a measure space. The dynamics of such automorphisms is studied within ergodic theory, nowadays a well-established branch of mathematics.

Now comes a crucial point: it turns out that Hamiltonian flows preserve a finer invariant, *a symplectic form*

$$\omega = \sum_{i=1}^{n} dp_i \wedge dq_i \tag{2}$$

on $\mathbb{R}^{2n}$. Note that the top wedge power $\omega^n = \omega \wedge \cdots \wedge \omega$ coincides with $n!\sigma$ and hence $h_t^*\omega = \omega$ yields the Liouville theorem.

Invariance of the symplectic form under Hamiltonian flows gave rise to a new mathematical discipline, symplectic topology. It deals with symplectic manifolds and their morphisms. Given a (necessarily even-dimensional) manifold M^{2n}, a symplectic form on M is a closed differential 2-form ω whose top power ω^n does not vanish and hence defines a volume form on M. At the first glance the symplectic form is yet another tensor field on M which should be studied along the lines of the standard differential (say, Riemannian) geometry. However, this is not the case: the classical Darboux theorem states that in appropriate local coordinates (p, q) near every point of M the form ω is given by (2). Thus local symplectic geometry does not exist: for instance there is no symplectic analogue of Riemannian curvature. The hunt for global invariants of subsets of symplectic manifolds became one of the central themes of modern symplectic topology. We shall return to this subject at the end of Sect. 2.4. Let us list some important examples of symplectic manifolds.

Example 2.1 A two-dimensional surface equipped with an area form is a symplectic manifold.

Example 2.2 Consider the cotangent bundle T^*X of a manifold X. Choose local coordinates $q_1, \ldots, q_n$ on X. Let $p_1, \ldots, p_n$ be the coordinates on the cotangent fibers T_q^*X associated with the basis $dq_1, \ldots, dq_n$. The 2-form ω given by formula (2) is symplectic. It does not depend on the specific choice of coordinates q and is called the standard symplectic form on T^*X.

Given a function $H : M \times \mathbb{R} \to \mathbb{R}$, the system (1) (understood in the Darboux coordinates (p, q)) gives rise to the Hamiltonian flow $h_t : M \to M$ which preserves the symplectic form ω. Individual diffeomorphisms h_t's obtained in this way are called *Hamiltonian diffeomorphisms*. They form a group denoted by $Ham\,(M, \omega)$. According to the ideology going back to Klein's program, in order to understand a geometric structure on a manifold, one should study its group of isometries. In our case this is the group $Symp(M, \omega)$ consisting of all *symplectomorphisms*, that is diffeomorphisms of M preserving the symplectic form M. In the case when the symplectic manifold M is closed and its first de Rham cohomology group vanishes, the identity component $Symp_0(M, \omega)$ coincides with $Ham\,(M, \omega)$. For general closed manifolds Ham is the commutator subgroup of $Symp_0$ and the quotient $Symp_0/Ham$ is "small". In this way the group of Hamiltonian diffeomorphisms appears as a central object of interest on the borderline between topology, geometry and dynamics (see [45]).

2.2 From Quasi-periodic Motion to Diffusion

The simplest (in terms of dynamical behavior) class of Hamiltonian flows on symplectic manifolds is given by *integrable systems*. In order to introduce this notion, recall that the Poisson bracket $\{F, G\}$ of functions F and G on a symplectic manifold M is given by the Lie derivative of F along the Hamiltonian flow generated by G. This operation is bilinear (over the reals) and anti-symmetric. In fact, it defines a natural Lie bracket on the space $C^\infty(M)/\{\text{constants}\}$ which can be canonically identified with the Lie algebra of the group $Ham\,(M, \omega)$.

We say that a Hamiltonian system associated with a time-independent Hamiltonian $H : M \to \mathbb{R}$ is *integrable* if there exist functions $H_1, \dots, H_n$ on M with $H_1 := H$ so that $\{H_i, H_j\} = 0$ for all i, j, and the union of the regular level sets of the *moment* map

$$\mathcal{H} : M \to \mathbb{R}^n, \qquad x \mapsto (H_1(x), \dots, H_n(x))$$

have full measure in M. Behavior of integrable systems is described by *the Liouville-Arnold theorem* (see [1, 12]). It states, in particular, that every compact connected regular level of the moment map, $\mathcal{H}^{-1}(c)$, is an n-dimensional torus, say L, which is invariant under the flow h_t and which carries a (quasi)-periodic motion. The latter means that in certain angular coordinates θ on $L = \mathbb{R}^n/\mathbb{Z}^n$ the dynamics looks like the shift $h_t\theta = \theta + tv$ for some $v \in \mathbb{R}^n$. The dynamics crucially depends on the arithmetic properties of the coordinates $v_1, \dots, v_n$ of the rotation vector v: For instance, if all of them are rational numbers, every trajectory on L is periodic, and if they are independent over $\mathbb{Q}$, every trajectory is uniformly distributed in L. In the latter case we shall call the vector v *incommensurable*.

Example 2.3 Let $M = T^*\mathbb{T}^n$ be the cotangent bundle of the n-dimensional torus equipped with coordinates $(p, q \mod \mathbb{Z}^n)$. In these coordinates the standard symplectic form is written as $\omega = \sum_{i=1}^n dp_i \wedge dq_i$ (see Example 2.2 above). In the language of classical mechanics this is the phase space of a system of pendulums consecutively connected to one another (see Wikipedia article “Double pendulum” for pictures and animations in the case $n = 2$). The Hamiltonian $H = H(p)$, which depends on the momenta variables only, is integrable near every regular energy level $\{H = \text{const}\}$: Indeed, $\{H, p_i\} = 0$ and $\{p_i, p_j\} = 0$ for all $i, j = 1, \dots, n$. Invariant tori carrying quasi-periodic motion are given by $\{p = p^0\}$, and the rotation vector on such a torus is $\frac{\partial H}{\partial p}(p^0)$. The significance of this example is due to a more advanced version of the Liouville-Arnold theorem: Near any invariant torus of a general integrable system one can choose coordinates (p, q) as above (called in this case *the action-angle coordinates*) so that in these coordinates the original Hamiltonian has the form $H = H(p)$.

Next, we address the question about the behavior of an integrable system under small perturbations of the Hamiltonian function. The *Kolmogorov–Arnold–Moser (KAM) theory* [2, 27, 42, 47] guarantees that (under certain non-degeneracy assumptions) the invariant n-dimensional tori carrying quasi-periodic motion persist provided the rotation vector v is “strongly” incommensurable, that is its coordinates $\{v_i\}$

do not admit anomalously small linear combinations with integer coefficients. Since incommensurable v's form a set of full measure, most of the tori persist. By persistence we mean that a torus becomes slightly deformed, it remains invariant under the perturbed Hamiltonian flow and carries the quasi-periodic motion with the same rotation vector v. This result was outlined by Kolmogorov in a short note of 1954, while the complete proofs were obtained by Arnold and Moser under various assumptions on the Hamiltonian. The KAM theory remains one of the finest analytical pieces of mathematical formalism of classical mechanics. From the viewpoint of functional analysis, the KAM-theorem lies in the framework of the Nash–Moser implicit function theorem in graded Fréchet spaces [55].

Suppose now that the symplectic manifold in question is four-dimensional (that is, in physicists' slang, our mechanical system has two degrees of freedom). By the energy conservation law, the motion of the system takes place in a three-dimensional energy level. Two-dimensional invariant tori described above divide the energy level into small invariant annuli of the form $\mathbb{T}^2 \times [0;1]$. These annuli serve as traps for the particle and, in particular, provide an obstruction to ergodicity of the system. This topological obstruction disappears for systems with $n \geq 3$ degrees of freedom. In his 1966 ICM talk [7] Arnold conjectured that in this case *generic* arbitrarily small perturbations $H(p,q,\epsilon) = H(p) + \epsilon H_1(p,q,\epsilon)$ of an integrable Hamiltonian $H = H(p)$ (in the notation of Example 2.3 above) admit trajectories $(p(t), q(t))$ which make long excursions through the gaps between KAM-tori: More precisely, there exists $T > 0$ so that $|p(T) - p(0)| > A$ for some positive constant A independent of the size of the perturbation. This phenomenon is called *Arnold diffusion*. It turnes out that even exhibiting *specific* (let alone generic) examples of this kind is a very hard problem. The first one was discovered by Arnold himself [4, 22]. The study of Arnold diffusion remains a popular theme in Hamiltonian dynamics. We refer to a beautiful paper by Kaloshin and Levi [41] for a more detailed survey.

2.3 Topology of Lagrangian Submanifolds

Invariant tori of integrable and near-integrable systems carrying quasi-periodic motion have a remarkable topological property: they are Lagrangian with respect to the symplectic structure. By definition, an n-dimensional submanifold L of a $2n$-dimensional symplectic manifold (M,ω) is *Lagrangian* if ω vanishes on each tangent space T_xL. In addition to the above-mentioned appearance in classical mechanics, Lagrangian submanifolds arise on various occasions in topology, algebraic geometry and the calculus of variations (see Sect. 4 for an example arising in the theory of wave propagation). The following basic examples will be important for the purposes of our exposition.

Example 2.4 A curve on a surface equipped with an area form is Lagrangian.

Example 2.5 Let f be a symplectomorphism of a symplectic manifold (M,ω). Then $\mathrm{graph}(f) \subset M \times M$ is Lagrangian with respect to the symplectic form $\omega \oplus -\omega$.

Example 2.6 Consider the cotangent bundle T^*X of a manifold X equipped with the standard symplectic form. A section of T^*X is Lagrangian if and only if it is the graph of a *closed* 1-form on X.

In the light of Example 2.4, multi-dimensional Lagrangian submanifolds can be considered as generalizations of curves on surfaces. Recall that given a closed immersed plane curve $\gamma : S^1 \to \mathbb{R}^2$, one can associate to it a topological invariant, the turning number. It is defined as the degree of the Gauss map $S^1 \to \mathbb{R}P^1$ which sends the point $t \in S^1$ to the tangent line $\mathbb{R} \cdot \dot{\gamma}(t)$ at the point $\gamma(t)$. Arnold discovered [6] that this construction admits a far-reaching generalization to Lagrangian immersions of the standard symplectic vector space $(\mathbb{R}^{2n}, \omega = dp \wedge dq)$ and that the Lagrangian analogue of the turning number is in fact an index introduced in Maslov's earlier work on quasi-classical approximation (see [44] and references therein).

Let us sketch Arnold's very elegant construction. Denote by Λ_n the Grassmannian of all linear Lagrangian subspaces of $\mathbb{R}^{2n} = \mathbb{C}^n$. The unitary group acts transitively on Λ_n, and the stabilizer of any given Lagrangian subspace coincides with the orthogonal group $O(n)$. Thus $\Lambda_n = U(n)/O(n)$, and hence we have a well-defined map

$$\phi : \Lambda_n \to S^1 \subset \mathbb{C}, \qquad [A] \mapsto (\det A)^2.$$

Here A stands for a matrix from $U(n)$ and $[A]$ for its equivalence class representing a Lagrangian subspace from Λ_n.

Write θ for the polar angle on the circle S^1. The cohomology class

$$\mu := \frac{1}{2\pi}\phi^*[d\theta] \in H^1(\Lambda_n; \mathbb{Z})$$

is called *the universal Maslov class*. Given a Lagrangian immersion $\gamma : L \to \mathbb{R}^{2n}$, consider the Gauss map $g : L \to \Lambda_n$ which takes a point $x \in L$ to the Lagrangian tangent subspace $\gamma_*(T_x L)$. The pull-back $\mu_L := g^*\mu \in H^1(L; \mathbb{Z})$ of the universal Maslov class is called the *Maslov class* of L. It measures the winding number of tangent planes to $\gamma(L)$ along 1-cycles in L.

This construction can be generalized, after some elementary topological considerations, to Lagrangian submanifolds in arbitrary symplectic manifolds. Due to Arnold's work the Maslov class entered the toolbox of the symplectic geometer. This notion plays a fundamental role in various modern developments.

Another invariant associated to a Lagrangian submanifold $L \subset \mathbb{R}^{2n}$ is the Liouville class $\lambda_L \in H^1(L; \mathbb{R})$. Its value on a 1-cycle $C \subset L$ equals $\int_C p\,dq$. When C is a circle, this is just the symplectic area of any disc in $\mathbb{R}^{2n}$ spanning C. We shall refer to the Maslov and the Liouville classes of L as the *classical invariants* of L.

The next definition will be important for our further discussion. A Lagrangian submanifold $L \subset \mathbb{R}^{2n}$ is called *monotone* if its Maslov and Liouville classes coincide: $\mu_L = \lambda_L$. An example is given by the split n-dimensional torus

$$L_0 = S^1(r) \times \cdots \times S^1(r) \subset \mathbb{R}^2 \times \cdots \times \mathbb{R}^2, \tag{3}$$

where $\pi r^2 = 2$.

The Maslov class of embedded (as opposed to immersed) Lagrangian submanifolds inherits certain rigidity properties from the turning number of simple closed curves in the plane which attains values ± 2 only. It was recently proved by Buhovsky [28] by Floer-homological methods that for every *monotone* Lagrangian torus L in

$\mathbb{R}^{2n}$ the image of $H_1(L;\mathbb{Z}) = \mathbb{Z}^n$ under the Maslov class μ_L equals $2\mathbb{Z}$. At the same time certain basic properties of the Maslov class are still far from being understood. For instance, it is unknown whether there exists a closed Lagrangian submanifold of $\mathbb{R}^{2n}$ whose Maslov class vanishes.

Let us focus now on the *Lagrangian knots problem* posed by Arnold in [15]. Nowadays this problem remains a very active research area. It can be informally formulated as follows: Consider the space of all Lagrangian submanifolds of (M,ω) with the same classical invariants. What are its connected components? Let us elaborate this question for monotone Lagrangian tori where the precise formulation is especially transparent. Lagrangian knots are simply the connected components of the space of all monotone Lagrangian tori. The split torus L_0 given by (3) plays the role of the trivial knot. The existence of non-trivial Lagrangian knots in this setting was discovered by Arnold's student Chekanov [31]. The complete classification is not yet understood even in the smallest non-trivial dimension $2n = 4$.

2.4 Lagrangian Intersections and Symplectic Fixed Points

Here we discuss Arnold's famous conjectures [3, 7, 12] which he discovered while analyzing the proof of Poincaré's "last geometric theorem" on fixed points of area-preserving maps of the annulus.

Consider the cotangent bundle T^*X of a closed manifold X. We equip it with the standard symplectic form and identify X with the zero section. We start our discussion with the following observation due to Arnold. Let ϕ be a Hamiltonian diffeomorphism of T^*X which is C^1-close to the identity. Then $\phi(X)$ is C^1-close to X, and hence is a section of the cotangent bundle. By Example 2.6 above, $\phi(X)$ is the graph of a closed 1-form, say α, on X. Furthermore, the condition that ϕ is Hamiltonian (and not just symplectic) translates into the fact that α is exact: $\alpha = dF$, where F is a smooth function on X. The critical points of F (that is the zeros of α) are in one-to-one correspondence with the intersection points $\phi(X)\cap X$. This consideration led Arnold to the following conjecture. Denote by $c(X)$ the minimal number of critical points of a smooth function on X.

Arnold's Lagrangian Intersection Conjecture The number of intersection points satisfies $|\phi(X)\cap X| \geq c(X)$ for *every* Hamiltonian diffeomorphism ϕ of T^*X.

To the best of our knowledge, the conjecture is still open as stated. Various partial results are known. In particular, when $\phi(X)$ is transverse to X, the number of intersection points is not less than the sum of Betti numbers of X [39, 43]. Let us emphasize that modern proofs are far-reaching generalizations of Arnold's Morse-theoretical argument.

The Weinstein normal form theorem [53] states that a tubular neighborhood of any closed Lagrangian submanifold X of an arbitrary symplectic manifold (M,ω) is symplectomorphic to a neighborhood of X in the cotangent bundle T^*X. Thus for any C^1-small Hamiltonian diffeomorphism ϕ of (M,ω) one has

$$|\phi(X)\cap X| \geq c(X). \qquad (4)$$

For certain classes of Lagrangian submanifolds this inequality (or its weaker version) extends to all, not necessarily C^1-small Hamiltonian diffeomorphisms. Take for instance a two-dimensional sphere S^2 equipped with the standard area form. Every small circle $X \subset S^2$ can be displaced by a suitable rotation and hence (4) is obviously violated. However when X is an equator (that is a simple closed curve dividing the area of the sphere into two discs of equal areas) (4) holds true by an obvious area control. This toy example shows that the extension of Arnold's Lagrangian intersection conjecture to arbitrary Lagrangian submanifolds is a delicate task even on the conjectural level, let alone proofs. Let us discuss one case of major importance which was discovered by Arnold himself.

Arnold's Fixed Points Conjecture Any Hamiltonian diffeomorphism f of a closed symplectic manifold (M, ω) has at least $c(M)$ fixed points.

Observe that fixed points of f are in one-to-one correspondence with the intersection points of graph(f) with the diagonal $\Delta \subset M \times M$. Both graph$(f)$ and Δ are Lagrangian submanifolds of $(M \times M, \omega \oplus -\omega)$ (cf. Example 2.5 above), and moreover graph(f) is the image of Δ under the Hamiltonian diffeomorphism $\mathbb{1} \times f$ of $M \times M$. Thus Arnold's fixed points conjecture can be reduced to the generalized Lagrangian intersection conjecture. Its current status is similar to the one of the Lagrangian intersection conjecture: it is open as stated, but numerous weaker statements are known starting from the pioneering work [32] by Conley and Zehnder (see e.g. [45, 46] for a more detailed account).

Arnold's conjectures served as a main motivation for development of some major techniques in modern symplectic topology such as the theory of generating functions and Floer theory. Let us discuss the latter very briefly (see [34, 35, 46] for further details). Let (M, ω) be a closed symplectic manifold. For the sake of simplicity we assume that $\pi_2(M) = 0$ (think for instance about the 2-torus). Consider the space $\mathcal{L}$ of free contractible loops $x : S^1 \to M$, where $S^1 = \mathbb{R}/\mathbb{Z}$. Take any time-periodic Hamiltonian function $H : M \times S^1 \to \mathbb{R}$. Define an *action functional*

$$\mathcal{A}_H : \mathcal{L} \to \mathbb{R}, \qquad x \mapsto \int_{S^1} H(x(t), t)\, dt - \int_D \omega,$$

where D is any disc spanning x. The topological condition $\pi_2(M) = 0$ guarantees that any two such discs are homotopic with fixed boundary and thus $\int_D \omega$ does not depend on the particular choice of the disc D. A version of the least action principle in classical mechanics states that the critical points of the functional $\mathcal{A}_H$ are in one-to-one correspondence with contractible 1-periodic orbits of the Hamiltonian flow h_t generated by H. Any such orbit, in turn, corresponds to a fixed point of the time-one-map $\phi_H := h_1$ of the Hamiltonian flow h_t. With this language Floer theory is a Morse theory (cf. [49]) for the action functional $\mathcal{A}_H$ on the space $\mathcal{L}$. A systematic development of such a Morse theory faces many difficulties.

The first difficulty is that the gradient flow of $\mathcal{A}_H$ is not defined in any reasonable sense. Fortunately, the trajectories of the gradient flow connecting critical points correspond to a well-posed Fredholm problem with asymptotic boundary conditions. Topologically they are cylinders (paths in the loop space $\mathcal{L}$) satisfying, after a suitable choice of a metric on $\mathcal{L}$, a version of the Cauchy-Riemann equations (here Floer theory meets Gromov's pseudo-holomorphic curves in symplectic manifolds). From the analytic viewpoint, these connecting trajectories solve an elliptic PDE. A Fredholm nature of this PDE guarantees that generically the spaces of connecting trajectories are finite-dimensional manifolds with "nice" compactifications.

The second difficulty is that the Morse indices of critical points of the action functional are infinite. Nevertheless, the index difference can be defined. Arnold's work on the Maslov index plays a crucial role in this construction.

A Morse-type theory built (starting from a work by Floer) along these lines yields topological lower bounds on the number of critical points of $\mathcal{A}_H$, and hence on the number of fixed points of the Hamiltonian diffeomorphism ϕ_H generated by H.

From the above discussion the reader might get the impression that the interaction between symplectic topology and Hamiltonian dynamics goes in the direction from topology to dynamics: a powerful machine of infinite-dimensional Morse theory on loop spaces solves a purely dynamical question on symplectic fixed points. In fact the interaction goes the other way round as well in the most fruitful way. In particular, it leads to new symplectic invariants of open domains in the standard symplectic vector space $\mathbb{R}^{2n}$. Given a domain $U \subset \mathbb{R}^{2n}$, denote by $\mathcal{H}_a$, $a > 0$, the set of all compactly supported time-independent Hamiltonian functions $H : U \to [0, a]$ which attain the maximal value a on a non-empty open subset of U. A beautiful phenomenon discovered by Hofer and Zehnder [40, 45] by methods of infinite-dimensional calculus of variations is as follows: there exists $A > 0$ so that for every $H \in \mathcal{H}_a$, $a \geq A$ the Hamiltonian flow of H possesses a *non-constant* closed orbit of period ≤ 1. Observe that constant closed orbits always exist and correspond to the critical points of H. By definition, the *Hofer–Zehnder capacity* $c_{\mathrm{HZ}}(U)$ is the infimum of such A. This capacity is invariant under symplectomorphisms and monotone under inclusions. In particular, if U admits a symplectic embedding to V one has $c_{\mathrm{HZ}}(U) \leq c_{\mathrm{HZ}}(V)$.

Denote by $B^{2n}(r)$ the standard Euclidean ball of radius r in $\mathbb{R}^{2n}$. A remarkable feature of the Hofer–Zehnder capacity is that it equals πr^2 both for the ball $B^{2n}(r)$ and for the cylinder $B^2(r) \times \mathbb{R}^{2n-2}$. In particular, the ball $B^{2n}(R)$ with $R > r$ does not admit a symplectic embedding into the cylinder $B^2(r) \times \mathbb{R}^{2n-2}$, even though the latter has infinite volume while the former has finite volume. This statement, known as Gromov's non-squeezing theorem (see [45]) became a logo of modern symplectic topology. The Hofer–Zehnder approach to its proof highlights the role of periodic orbits of Hamiltonian flows envisioned by Arnold.

3 Singularity Theory

In the seventies I started most of my papers with the words: 'there exists an interesting and unexpected relation between...' (*the continuations being different in the different papers*).

Arnold's speech on 20 June 1994 when receiving the Universidad Complutense Honoris Causa degree in Sciences, Spain

We review a part of singularity theory concerned with isolated singularities of analytic functions. We describe mostly the holomorphic case; the results carry over to the real analytic case with minor changes which we will indicate.

Our aim is to explain a remarkable discovery of Arnold from the early 1970's: simple singularities are classified by the $A - D - E$ Dynkin diagrams, [11]. This discovery, connecting singularities with Lie algebras, reflection groups and invariant theory, has put a new face on singularity theory and stimulated extensive and intensive developments both in the theory and in its applications to variational problems.

For the singularities of maps and of complete intersections, for global singularity theory, the theory of non-isolated singularities, and for a lot of interesting and important applications (including Legendrian singularities, bifurcations in dynamical systems, singularities of boundaries of functional domains, asymptotics of oscillating integrals, etc.) we refer the reader to the two volume edition on singularity theory and its applications [23, 24].

3.1 Main Notions

3.1.1 Singularities

Singularity theory studies the local behavior of a function near a *critical point*, i.e., a point where the first differential of the function vanishes. The value of a function at a critical point is called a *critical value*.

We encounter a baby version of singularity theory already in high school calculus, in the context of real-valued functions. We learn there that if the first derivative $f'(x)$ of a function f vanishes at some point x_0, while $f''(x_0) \neq 0$, the point x_0 is either a local minimum or a local maximum. If $f'(x_0) = f''(x_0) = 0$, while the third derivative does not vanish, x_0 is an inflection point. In the case of "higher order" singularities and of functions of several variables the story becomes much more complicated and new tools and ideas are required. This is what singularity theory is about. In what follows we stick to the case of *holomorphic* functions of several complex variables; this case is somewhat simpler and more transparent.

We may and will assume that the critical point under consideration is the origin $O \in \mathbb{C}^n$ and that $f(O) = 0$. Denote by $\mathcal{O}_n$ the ring of the holomorphic function germs $f : \mathbb{C}^n \to \mathbb{C}$ at the origin $O \in \mathbb{C}^n$, and by $\mathfrak{m}_n \subset \mathcal{O}_n$ the maximal ideal, i.e., the subring of germs vanishing at the origin, $f(O) = 0$. Unless otherwise stated we shall always deal with *germs* of functions, manifolds etc. near the origin.

The group $\mathcal{D}_n$ of biholomorphic maps $g : (\mathbb{C}^n, O) \to (\mathbb{C}^n, O)$ acts on $\mathcal{O}_n$ as follows,

$$g(f) = f \circ g^{-1}, \quad \forall f \in \mathcal{O}_n, \quad \forall g \in \mathcal{D}_n.$$

Two functions $f_1, f_2 \in \mathfrak{m}_n$ are *equivalent*, $f_1 \sim f_2$, if they belong to the same orbit of the action; the equivalence class, i.e., the orbit of the action, is a *singularity*.

By Arnold, singularity theory, like life sciences, is subdivided into zoology, anatomy and physiology. The "zoology of singularities" is the description of what, where and how singularities can be encountered; the "anatomy and physiology of singularities" studies their structure and how they function.

3.1.2 The Morse Lemma

A critical point is *non-degenerate* if the second differential of the function at this point is a non-degenerate quadratic form. *The Morse lemma* says that a holomorphic function $f \in \mathfrak{m}_n$ near a non-degenerate critical point is equivalent to a non-degenerate quadratic form,

$$f \sim x_1^2 + \cdots + x_n^2.$$

Any small deformation of a non-degenerate critical point has the same singularity close to the origin. This is the simplest singularity.[6]

There is a natural way to define an equivalence relation for critical points of functions of different number of variables. Namely, holomorphic functions $f \in \mathfrak{m}_n$ and $h \in \mathfrak{m}_m$ are *stably equivalent* if they become equivalent after adding the squares of supplementary variables:

$$f(x_1, \ldots, x_n) + x_{n+1}^2 + \cdots + x_k^2 \sim h(y_1, \ldots, y_m) + y_{m+1}^2 + \cdots + y_k^2.$$

According to a theorem of A. Weinstein (1971), two functions of the same number of variables are stably equivalent if and only if they are equivalent; thus one can simultaneously study the singularities of functions of different number of variables. In particular, all non-degenerate critical points are stably equivalent to x^2. This is the singularity the classification begins with. It is called the *Morse singularity* and is denoted by A_1.

3.1.3 Multiplicity

Under a small deformation, a degenerate isolated critical point decomposes into a finite number of "simpler" ones. If a small deformation is *generic*, then all the "simpler" singularities are non-degenerate. Their number is called the *geometric multiplicity* of an isolated singularity. As Fig. 1 illustrates, the geometric multiplicity of x^2 is 1, and of x^3 is 2.

We use the notation of Sect. 3.1.1. First of all, let us introduce an important algebraic invariant of a singularity f: The ideal $\mathcal{I}_f \subset \mathcal{O}_n$ generated by the partial derivatives of $f \in \mathfrak{m}_n$ is called the *gradient ideal* of f, and the quotient $\mathcal{Q}_f = \mathcal{O}_n/\mathcal{I}_f$ is the *local algebra* of f.

The dimension of the local algebra over $\mathbb{C}$ is the *algebraic multiplicity* of f. A classical result (whose complete proof was published by V. Palamodov in 1967) asserts that the geometric and the algebraic multiplicities of an isolated critical point coincide.[7] That is, under a generic small deformation, an isolated critical point of f decomposes into $\mu(f) = \dim_{\mathbb{C}} \mathcal{Q}_f$ critical points of type A_1.

[6]Over $\mathbb{R}$, we have more possibilities: $f \sim x_1^2 \pm \cdots \pm x_n^2$. The real analytic case always needs such an adjustment; in what follows we leave it to reader.

[7]In fact, a critical point of a holomorphic (or real analytic) function is isolated if and only if its multiplicity is finite, [23, 24].

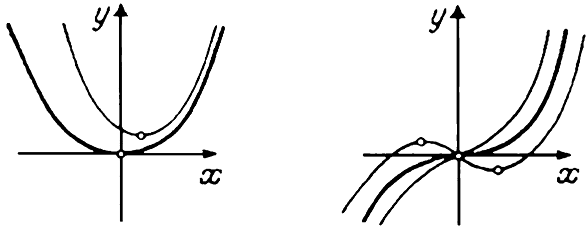


Fig. 1 Small deformations of x^2 and of x^3

Example 3.1 (1) For A_1-singularity, we have $(x^2)' = 2x$, hence the ideal $\mathcal{I}_{A_1}$ is generated by $2x$ and the quotient $\mathcal{Q}_{A_1}$ by 1, that is, $\mathcal{Q}_{A_1} \cong \mathbb{C}$ is one-dimensional, and (non-surprisingly!) $\mu(A_1) = 1$.

(2) The function x^3 (as well as $x^3 + Q(y)$, where $Q(y)$ is a non-degenerate quadratic form, e.g., $Q(y) = y_1^2 + \cdots + y_k^2$) has the simplest degenerate critical point at 0, denoted by A_2. The ideal $\mathcal{I}_{A_2}$ is spanned by x^2 (by $x^2, y_1, \ldots, y_k$, resp.), the local algebra $\mathcal{Q}_{A_2}$ is generated by 1 and x over $\mathbb{C}$, and $\mu(A_2) = 2$.

(k) Similarly, A_k-singularity is given by x^{k+1} and $\mu(A_k) = k$.

According to a theorem of J.-C. Tougeron (1968), any holomorphic function at a critical point of multiplicity μ is equivalent to its Taylor polynomial of degree $\mu + 1$; in particular, every singularity of finite multiplicity has a polynomial representative.

3.1.4 Versal Deformation

According to Example 3.1(2), under a small deformation the singularity A_2 decomposes into two non-degenerate ones. For instance, one can take as a deformation of x^3 (of $x^3 + Q(y)$, resp.) the function $x^3 + \epsilon x$ (or $x^3 + Q(y) + \epsilon x$, resp.) which has two non-degenerate critical points, $x_{1,2} = \pm\sqrt{-\epsilon}$ (and $y_1 = \cdots = y_k = 0$, resp.) near the origin, see Fig. 1.

This consideration shows that A_2 is a *non-generic* singularity of a single holomorphic function: it is destroyed by a small perturbation.

However, when one investigates a *family* of holomorphic functions, e.g., $f_t(x) = x^3 + tx$, every nearby family does have the singularity A_2 for some value of the parameter t close to 0. Thus a non-generic singularity of an individual function becomes a generic singularity of a family.

This discussion leads us to the study of families in which a given singularity appears as a generic one. In particular, one has to understand *bifurcations* of the singularity, that is the ways it decomposes into simpler ones under small changes of the parameters of the family. We shall refer to a family as to a *deformation* of the singularity. Remarkably, if a singularity is isolated, then it is sufficient to study only one, so-called *versal*, deformation of a function $f \in \mathfrak{m}_n$ which is transversal to the orbit of f under the action of $\mathcal{D}_n$: it turns out that every deformation of f can be induced (in some natural way) from the versal one.

In particular, the family

$$F(x, \lambda) = f(x) + \sum_{j=1}^{\mu} \lambda_j \phi_j(x), \tag{5}$$

where μ is the multiplicity of f and $\phi_j(x)$'s are representatives of a basis of the local ring $\mathcal{Q}_f$, provides a versal deformation.

For example, a versal deformation of the A_k-singularity given by x^{k+1} is

$$F(x,\lambda) = x^{k+1} + \sum_{j=1}^{k} \lambda_j x^{k-j}. \tag{6}$$

3.1.5 Bifurcation Sets

Let us pass to a topological analysis of singularities. An important character of our story is a hypersurface in the base of a versal deformation called the *discriminant* (or the *level bifurcation set*) of a singularity. Fix $f \in \mathfrak{m}_n$ having a critical point at O of multiplicity μ and its versal deformation $F(x,\lambda)$. By Sard's lemma, near the origin the level set

$$V_\lambda = \{x : F(x,\lambda) = 0\}$$

is a smooth manifold for almost all values of the parameter λ. The values of λ for which V_λ is *singular* form the *discriminant* (or *level bifurcation set*) Σ_f of f. This hypersurface appears in various applications of singularity theory (we discuss an example of this kind in Sect. 4).

Example 3.2 (1) For the function $x^3 + y^2$ (the A_2-singularity) and its versal deformation $F(x,y,\lambda) = x^3 + y^2 + \lambda_1 x + \lambda_2$, the discriminant $\Sigma_{A_2} \subset \mathbb{C}^2_{\lambda_1,\lambda_2}$ is formed by those $\lambda = (\lambda_1, \lambda_2)$ such that the level curve

$$V_\lambda = \{(x,y) : F(x,y,\lambda) = 0\} \subset \mathbb{C}^2_{x,y}$$

is singular. That is, there exists a point in V_λ where the gradient of $F(x,y,\lambda)$ vanishes, $F'_x = F'_y = 0$, i.e., the system $x^3 + y^2 + \lambda_1 x + \lambda_2 = 3x^2 + \lambda_1 = y = 0$ has a solution. We get

$$\Sigma_{A_2} = \{\lambda_1 = -3x^2,\ \lambda_2 = 2x^3, x \in \mathbb{C}\} = \{4\lambda_1^3 + 27\lambda_2^2 = 0\} \subset \mathbb{C}^2_{\lambda_1,\lambda_2}.$$

This curve is called a *cusp*.

(2) Similarly, for the A_k-singularity given by $x^{k+1} + y^2$, the discriminant Σ_{A_k} is formed by the values of λ's such that the polynomial $F(x,y,\lambda) = x^{k+1} + y^2 + \sum_{j=1}^{k} \lambda_j x^{k-j}$ has a critical point with zero critical value. The surface $\Sigma_{A_k} \subset \mathbb{C}^k$ is called the (generalized, if $k > 3$) *swallowtail surface*.

Real parts of Σ_{A_2} and Σ_{A_3} are shown on Figs. 2 and 3.

In some problems one should deal with deformations of an isolated singularity inside $\mathfrak{m}_n$, as opposed to $\mathcal{O}_n$ (e.g. in the study of Lagrangian singularities, see Sect. 4 below): In other words we impose a constraint $F(O,\lambda) \equiv 0$. In this context the deformation transversal to the orbit of the singularity is called a *truncated* versal deformation. It depends on $\mu - 1$ parameters, where μ is the multiplicity of the singularity. The *caustic* or the *function bifurcation set* of an isolated singularity is a hypersurface

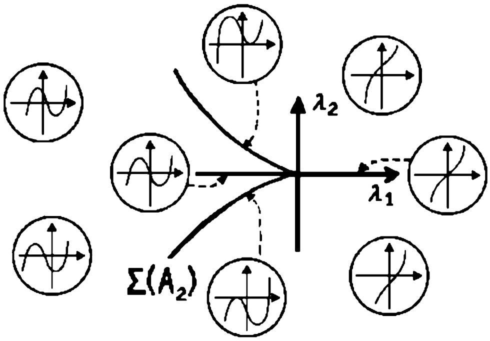


Fig. 2 Typical level sets and the discriminant of A_2-singularity

formed by the parameters of a truncated versal deformation that correspond to the functions having a *degenerate* critical point. The caustic can be described also as the set of the critical values of a projection of the discriminant to the base of the truncated versal deformation along a *generic* direction.

Continuation of Example 3.2 One obtains a truncated versal deformation of A_k singularity from the versal deformation (6) by setting $\lambda_k = 0$. A generic direction in the space of versal deformation is transversal to the swallowtail surface Σ_k at the origin. An easy exercise shows that the function bifurcation set of A_k is diffeomorphic to $\Sigma_{A_{k-1}}$.

3.1.6 Digression on the Real Case: Perestroikas

In the real case, the discriminant, being a hypersurface in $\mathbb{R}^\mu$, divides the base of the versal deformation into domains. For values of λ in the same domain, the level sets have the same shape. When λ goes trough the discriminant, the level set changes its shape, i.e., a metamorphosis, or a *perestroika*[8] occurs.

End of Example 3.2 For the versal deformation of the real function x^3, parameters λ's inside the cusp give polynomials with three real roots, whereas outside with only one real root. Parameters on the cusp (not at the vertex) give polynomials with one multiple root (A_1-singularity), and the vertex corresponds to x^3 (A_2-singularity), see Fig. 2.

For $k = 3$, the swallowtail surface[9] is shown on Fig. 3, together with its plane sections.

3.1.7 Monodromy Group

We keep notation of Sects. 3.1.4 and 3.1.5. Let $f(x)$ have an isolated critical point of multiplicity μ at $O \in \mathbb{C}^n$, and let $F(x, \lambda)$ be its versal deformation, e.g. as in (5). The

[8] Arnold was proud he had introduced this Russian word—it was extremely popular in the Soviet Union of the late 80's—in the international mathematical terminology.

[9] The swallowtail is one of the protagonists in singularity theory. It appears in different problems and in different contexts. The second named author, who attended Arnold's Seminar during almost 20 years starting from 1975, does not remember a single seminar with no swallowtail surface drawn on the blackboard…

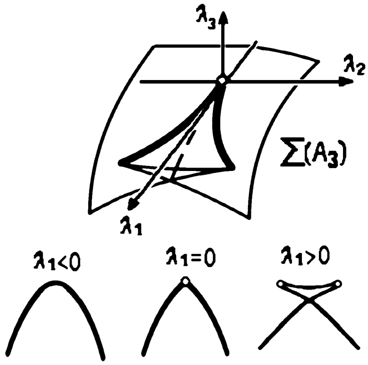


Fig. 3 The level bifurcation set of A_3-singularity and its plane sections

discriminant $\Sigma_f \subset \mathbb{C}^\mu$, as a complex hypersurface, has real codimension 2 and does not divide the base of a versal deformation. Therefore near the origin the non-singular level sets V_λ's all have the same topological type. By a theorem of Milnor (1968) they are homotopy equivalent to a wedge (bouquet) of μ spheres of real dimension $n-1$.

Fix a non-singular parameter value λ_* and abbreviate $V_* := V_{\lambda_*}$. The only non-trivial integer homology is $H_{n-1}(V_*) = \mathbb{Z}^\mu$.

Consider the union of the level sets V_λ for all λ's near the origin. This is a hypersurface, $V_\Lambda = \{F(x,\lambda) = 0\} \subset \mathbb{C}^{n+\mu}$.

Denote by $\widehat{\Lambda} \subset \mathbb{C}^\mu$ the complement of the discriminant Σ_f, and by $\widehat{V}_\Lambda$ the preimage of $\widehat{\Lambda}$ under the canonical projection $(x,\lambda) \mapsto \lambda$. We get a locally trivial fibration $\widehat{V}_\Lambda \to \widehat{\Lambda}$ which is called the *Milnor fibration* of f. The homology bundle associated with this fibration defines in a natural way a representation of the fundamental group $\pi_1(\widehat{\Lambda}, \lambda_*)$ in the integer homology of a non-singular level set,

$$\Gamma : \pi_1(\widehat{\Lambda}, \lambda_*) \to Aut\,(H_{n-1}(V_*))\,.$$

The image of this representation is the *monodromy group* of f. The monodromy group does not depend on a choice of a versal deformation, and is determined only by the type of the singularity.

3.1.8 Basis of Vanishing Cycles

In order to construct a basis in $H_{n-1}(V_*)$, we first consider the case $\mu = 1$, when the basis consists of one cycle only.

Example 3.3 If the critical point of f is non-degenerate, then in some local coordinate system we have

$$f(x) = x_1^2 + \cdots + x_n^2, \qquad F(x,\lambda) = x_1^2 + \cdots + x_n^2 - \lambda.$$

A non-singular level set V_* is diffeomorphic to the tangent bundle T_*S^{n-1}, where

$$S^{n-1} = \{x \in \mathbb{C}^n : \ x_1^2 + \cdots + x_n^2 = 1,\ \operatorname{Im} x_j = 0,\ j = 1, \ldots, n\}$$

is the standard $(n-1)$-dimensional unit sphere.

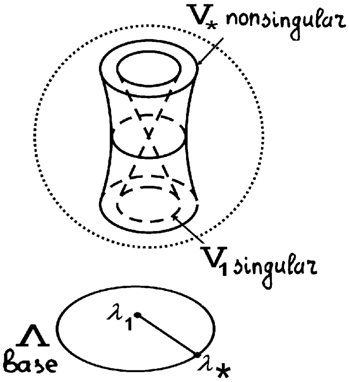


Fig. 4 Vanishing cycle, $\mu = 1$

Take the path $\phi(t) = 1 - t$ in the parameter space, where $t \in [0; 1]$ is real, from the non-singular $\lambda_* = \phi(0) = 1$ to the singular $\lambda_1 = \phi(1) = 0$. In the non-singular level sets $V_{\phi(t)}$, $0 \le t < 1$, the spheres $S_t = \sqrt{1-t}\,S^{n-1}$ appear. If an orientation of S^{n-1} is chosen, then the S_t's are oriented as well. The integer homology $H_{n-1}(V_*)$ is generated by the homology class $[S_0] = \Delta \in H_{n-1}(V_*)$, called a *vanishing cycle*, as S_t vanishes (degenerates to a point) at $t = 1$. The case $n = 2$ is shown on Fig. 4.

For a singularity f of multiplicity μ we select μ vanishing cycles in the following way. Take a *generic* complex line $\mathbb{C}^1$ in $\mathbb{C}^\mu$ through a fixed non-singular value λ_*. This complex line intersects Σ_f at μ different points, say, $\lambda^{(1)}, \dots, \lambda^{(\mu)}$. Each of $\lambda^{(j)}$ corresponds to a non-degenerate critical point. In the language of deformations, we choose a *generic one-parameter deformation* of f. Then for each of μ different values of the parameter, $\lambda^{(1)}, \dots, \lambda^{(\mu)}$, the corresponding level set has a non-degenerate singular point.

Let the points λ_* and $\lambda^{(j)}$ are located on the complex line $\mathbb{C}^1$ as shown on Fig. 5. Exactly as in Example 3.3, the fiber V_* contains an $(n-1)$-dimensional sphere which vanishes along the segment $[\lambda_*; \lambda^{(j)}]$. Denote by $\Delta_j \in H_{n-1}(V_*)$ its homology class. The μ loops of Fig. 5, originating at λ_*, generate $\pi_1(\mathbb{C}^1 \setminus (\mathbb{C}^1 \cap \Sigma_f), \lambda_*)$. Hence by a theorem of Zariski they generate also $\pi_1(\widehat{\Lambda}, \lambda_*)$.

Example 3.4 For the A_2-singularity x^3, take a one-parametric deformation $F(x, \lambda) = x^3 - 3x - \lambda$. Any non-singular level consists of three points, $V_* = \{x_1, x_2, x_3\}$. Assume that these points are real and $x_1 < x_2 < x_3$. Take $\lambda_* = 0$. Then $V_* = V_0 = \{-\sqrt{3}, 0, \sqrt{3}\}$. Exactly two values, $\lambda = \pm 2$, lie on the discriminant: $V_{\pm 2} = \{\mp 1, \pm 2\}$. The points $x = \mp 1 \in V_{\pm 2}$ are non-degenerate critical points, $x^3 - 3x \mp 2 = (x \pm 1)^2(x \mp 2)$. The cycle $\Delta_1 = [x_2] - [x_1]$ (resp. $\Delta_2 = [x_3] - [x_2]$) vanishes along the segment $[0; 2]$ (resp. $[0; -2]$). These two cycles form a basis in $H_0(V_*) \cong \mathbb{Z}^2$ (in the case $n = 1$ the homology are assumed to be reduced modulo a point). The monodromy along the loop going around $\lambda = 2$ permutes x_2 and x_1, and going around $\lambda = -2$ permutes x_3 and x_2, see Fig. 6.

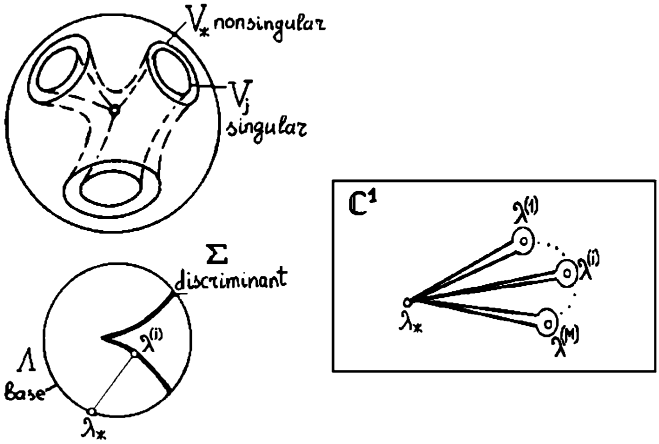


Fig. 5 Construction of a vanishing cycles basis

3.1.9 Intersection Matrix and Picard–Lefschetz Formula

The non-singular level set $V_* \subset \mathbb{C}^n$ is an oriented manifold of real dimension $(2n-2)$, therefore the intersection index $(\cdot \circ \cdot)$ in the integer homology $H_{n-1}(V_*) \cong \mathbb{Z}^\mu$ is a well-defined $\mathbb{Z}$-bilinear form. The matrix $(\Delta_i \circ \Delta_j)_{1\le i,j\le\mu}$ is called the *intersection matrix* of a singularity.

In particular, the self-intersection index of every vanishing cycle Δ_j is

$$(\Delta_j \circ \Delta_j) = (-1)^{(n-1)(n-2)/2}(1+(-1)^{n-1}). \tag{7}$$

It depends on n: it vanishes for even n, and it equals ± 2 for odd n.

The intersection matrix of a singularity changes under stabilization, however it changes in a predictable way. Namely, let $\tilde{f}$ be obtained from f by adding the squares of m additional variables. The loops γ_j's of Sect. 3.1.8 define also the vanishing cycles $\tilde{\Delta}_j$ for $\tilde{f}$, up to orientation. By a theorem of S. Gussein–Zade (1977), for a suitable choice of orientation,

$$(\tilde{\Delta}_i \circ \tilde{\Delta}_j) = (\mathrm{sign}(j-i))^m(-1)^{mn+m(m-1)/2}(\Delta_i \circ \Delta_j).$$

Therefore a class of stably equivalent singularities has exactly four distinct intersection matrices which can be reconstructed from one another. Two of them are symmetric, and two are skew-symmetric, differing by sign.

Continuation of Example 3.4 For the singularity A_2, it is easy to write down the four intersection matrices corresponding to the vanishing cycles Δ_1, Δ_2 and to $n \equiv 1, 2, 3, 0 \pmod 4$, respectively:

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

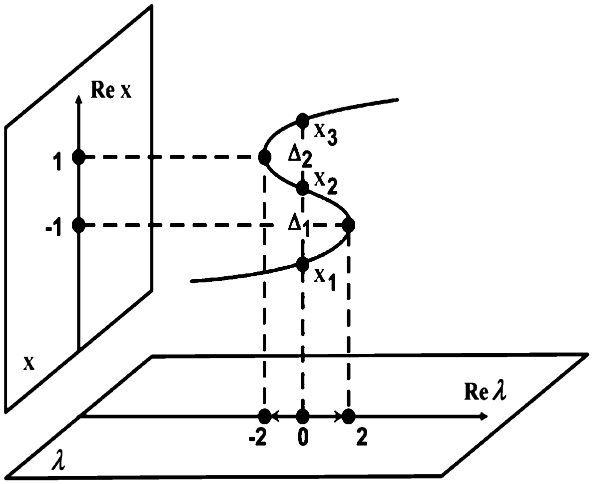


Fig. 6 Construction of a basis of vanishing cycles for A_2-singularity

The case of the A_k-singularity is similar. There are k vanishing cycles $\Delta_1, \ldots, \Delta_k$. The intersection matrix for $n \equiv 3 \pmod 4$ has (-2)'s on the diagonal, 1's on sub- and superdiagonals, while all other matrix elements vanish.

For a singularity of multiplicity μ, the monodromy group is generated by μ *Picard–Lefschetz operators*,

$$h_j : H_{n-1}(V_*) \to H_{n-1}(V_*), \quad 1 \le j \le \mu,$$

where h_j corresponds to the loop γ_j (or to the vanishing cycle Δ_j) described in Sect. 3.1.8. In terms of the intersection index,

$$h_j(\sigma) = \sigma + (-1)^{n(n+1)/2}(\sigma \circ \Delta_j)\Delta_j, \quad \forall \sigma \in H_{n-1}(V_*). \tag{8}$$

End of Example 3.4 For the A_2 singularity, we easily calculate the matrices of the operators h_1 and h_2 in the basis Δ_1, Δ_2,

$$[h_1] = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad [h_2] = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$

3.1.10 Dynkin Diagram of a Singularity

The intersection matrix of a class of stably equivalent singularities, in a chosen basis of vanishing cycles, can be described by a graph called the *Dynkin diagram of a singularity*. Take $n \equiv 3 \pmod 4$, so that the intersection matrix is symmetric, with (-2) on the diagonal. The vertices of the graph correspond to the vanishing cycles $\Delta_1, \ldots, \Delta_\mu$; the edge between Δ_i and Δ_j has multiplicity $|(\Delta_i \circ \Delta_j)|$; if the intersection index is negative, the corresponding edge is dotted.

The intersection matrix and the matrices of Picard–Lefschetz operators are easily reconstructed from the Dynkin diagram.

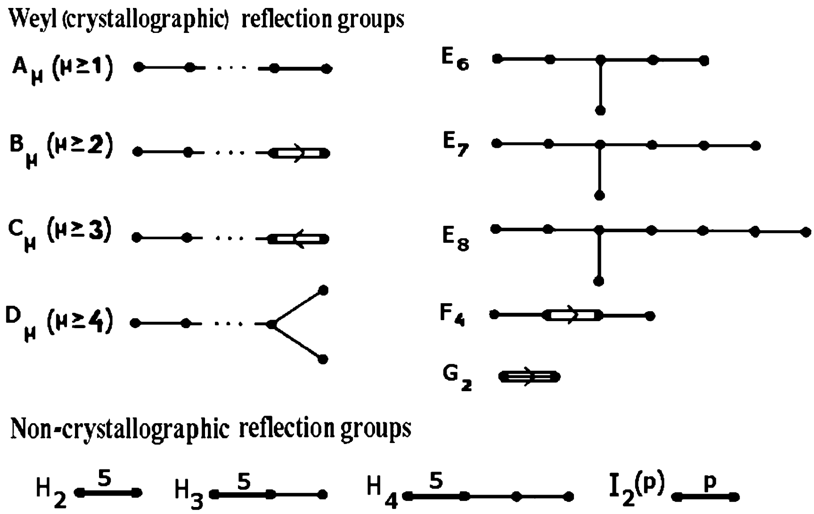


Fig. 7 Dynkin diagrams of reflection groups

In the case A_2 we get $\bullet - \bullet$. For the singularity A_k, the diagram contains k vertices:

$$A_k \qquad \bullet - \bullet - \cdots - \bullet - \bullet . \tag{9}$$

3.2 Simple Singularities and Reflection Groups

3.2.1 Simple Singularities

We use the notation of Sect. 3.1.1.

A holomorphic function $f \in \mathfrak{m}_n$ has a *simple* singularity at O, if a neighborhood of f in $\mathcal{O}_n$ is covered by a finite number of orbits. In other words, small deformations of a simple singularity give only a finite number of singularities.

The simple critical points of holomorphic functions are as follows

$A_k, k \geq 1$	$D_k, k \geq 4$	E_6	E_7	E_8
$x^{k+1} + y^2$	$x^2y + y^{k-1}$	$x^3 + y^4$	$x^3 + xy^3$	$x^3 + y^5$

(10)

The multiplicity is given by the subscript. If the number of variables is greater than two, one should add a non-degenerate quadratic form of the missing variables. It is an easy exercise to write down versal deformations. Dynkin diagrams of the simple singularities are presented on Fig. 7.

The interaction between the simple singularities of holomorphic functions and irreducible finite reflection groups goes in both directions. Starting with a reflection group, one can get a singularity, together with its versal deformation and the discriminant.

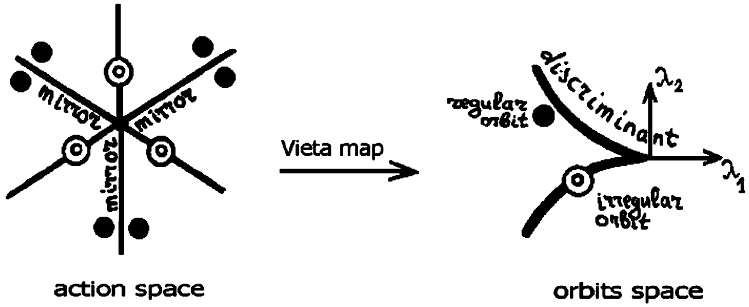


Fig. 8 Reflection group A_2

3.2.2 Arnold's First Example

We reproduce here Arnold's explanation of how the singularity A_2 appears in the study of the corresponding reflection group, [17, pages 60–61]. This was one of his favorite examples.

In the Euclidean plane $\mathbb{R}^2$, take three lines passing through the origin so that the angle between each pair of these lines equals $2\pi/3$. The group A_2 is generated by the reflections in these lines. It contains six elements, and a regular (generic) orbit consists of six distinct points, see Fig. 8.

Identify $\mathbb{R}^2$ with the plane $\Pi = \{x_1 + x_2 + x_3 = 0\}$ lying in the Euclidean space $\mathbb{R}^3$ equipped with coordinates x_1, x_2, x_3. The group A_2 acts on $\mathbb{R}^3$ by permutation of the coordinates. This action is generated by reflections in the mirrors $x_i = x_j$, $1 \le i \ne j \le 3$. Its restriction to the invariant plane Π coincides with the original action of A_2. The orbits in Π are unordered triples of real numbers with zero sum. The regular orbits are the triples of *different* numbers, and the irregular ones are the triples where two or three numbers coincide.

Now complexify the picture: the 3-dimensional space, the plane Π, the mirrors, the reflections and the group action. In other words, assume that $x_1, x_2, x_3 \in \mathbb{C}$ and keep all the formulas. The complexified orbit space consists of the unordered triples of complex numbers with zero sum. Treating an unordered triple as the roots of a cubic monic polynomial, $x^3 + \lambda_1 x + \lambda_2$, $(\lambda_1, \lambda_2) \in \mathbb{C}^2$ (the coefficient of x^2 vanishes, by the Vieta Theorem), we get a set of complex polynomials $\{x^3 + \lambda_1 x + \lambda_2\} \cong \mathbb{C}^2_{\lambda_1,\lambda_2}$.

The map $(x_1, x_2, x_3) \mapsto (\lambda_1, \lambda_2)$ from Π to the orbit space is the *Vieta map*. The orbit defined by $(\lambda_1, \lambda_2) \in \mathbb{C}^2$ is regular if and only if $x^3 + \lambda_1 x + \lambda_2$ has three distinct roots. The image of the mirrors under the Vieta map is the *IO (irregular orbits) variety* of the group.[10]

We see that the complexified orbit space is exactly the base of a versal deformation of the singularity A_2, and the IO variety is exactly the discriminant of the singularity A_2, cf. Example 3.2(1). The real part of the Vieta map is shown on Fig. 8.

[10] It is also called the *discriminant* of a reflection group; we follow Arnold's initial terminology (see e.g. [14]), in order not to confuse with the discriminant of a singularity.

Similarly, the group A_k acts on the hyperplane

$$\Pi_k = \{x_1 + \cdots + x_{k+1} = 0\} \subset \mathbb{R}^{k+1}$$

and is generated by reflections in the mirrors $x_i = x_j$, $1 \le i \ne j \le k$. After the complexification, the orbit space is identified with complex polynomials

$$\{x^{k+1} + \lambda_1 x^{k-1} + \cdots + \lambda_{k-1} x + \lambda_k\} \cong \mathbb{C}^k$$

which is the base of a versal deformation of the holomorphic singularity x^{k+1}, and the image of mirrors is the (generalized) swallowtail surface, i.e., the discriminant of the singularity A_k, cf. Example 3.2(2).

3.2.3 Root Systems

The group A_k is known as the *Weyl group* of the root system of the simple Lie algebra $sl(k+1)$. Recall some preliminaries [29].

A *root system* $\mathcal{R}$ is a finite system of non-zero vectors, called *roots*, in a finite dimensional Euclidean space $(E, (\cdot,\cdot))$ satisfying the following properties:

- the roots span E;
- if $\vec{v} \in \mathcal{R}$, then $-\vec{v} \in \mathcal{R}$ is the *only*[11] scalar multiple of $\vec{v}$ in $\mathcal{R}$;
- for every two roots $\vec{v}, \vec{u} \in \mathcal{R}$, the reflection of $\vec{u}$ in the hyperplane orthogonal to $\vec{v}$, is also a root:

$$\vec{u} - \frac{2(\vec{v},\vec{u})}{(\vec{v},\vec{v})}\vec{v} \in \mathcal{R}.$$

The reflections in the hyperplanes orthogonal to the roots generate the reflection group called the *Coxeter group* of the root system.

If in addition $2(\vec{v},\vec{u})/(\vec{v},\vec{v}) \in \mathbb{Z}$ for every two roots $\vec{v}, \vec{u} \in \mathcal{R}$, then the root system is called *crystallographic*. The corresponding reflection group preserves the integer lattice generated by the roots.

Any crystallographic root system has a *set of simple roots*: they form a basis in E, and each root with respect to this basis has either all non-negative or all non-positive integer coordinates.

Moreover, any pair of non-orthogonal simple roots $\vec{v}, \vec{u}$ is in one of the following positions:[12]

(1) they are of the same length and form an angle $\frac{2}{3}\pi$;
(2) the ratio of their lengths is $\sqrt{2}$, and they form an angle $\frac{3}{4}\pi$;
(3) the ratio of their lengths is $\sqrt{3}$, and they form an angle $\frac{5}{6}\pi$.

[11] This condition means that $\mathcal{R}$ is *reduced*; however, we do not consider non-reduced root systems. Also, we consider only *irreducible* root systems (that is, they can not be decomposed into a sum of root systems of smaller-dimensional spaces).

[12] This follows from the condition $2(\vec{v},\vec{u})/(\vec{v},\vec{v}) \in \mathbb{Z}$ and the fact that the angle between two non-orthogonal simple roots should be obtuse.

There is a one-to-one correspondence between the irreducible crystallographic root systems and the simple Lie algebras; the corresponding reflection groups are the Weyl groups of the simple Lie algebras, [29].

Initially, *Dynkin diagrams* were introduced to describe root systems. The vertices of the Dynkin diagram correspond to the simple roots, and the edge between two vertices has multiplicity 1, 2, or 3, according to the three possibilities (no edge between orthogonal roots). Edges of multiplicity 2 and 3 are oriented from the longer root to the shorter one.

Example 3.5 Root System A_k The roots of A_k are $e_i - e_j$, $1 \le i \ne j \le k+1$, where $\{e_1, e_2, \dots, e_{k+1}\}$ is an orthonormal basis in $\mathbb{R}^{k+1}$; all the roots have the same length $\sqrt{2}$; the simple roots are $e_{j+1} - e_j$, $1 \le j \le k$; simple roots $e_{j+1} - e_j$ and $e_{i+1} - e_i$ are orthogonal for $|i - j| > 1$. Thus the root system A_k has the same Dynkin diagram as the singularity A_k, see (9).

In particular, for A_2 the are two simple roots: $e_2 - e_1$ and $e_3 - e_2$. They correspond to the vanishing cycles Δ_1, Δ_2 of Example 3.4, see Fig. 6.

The irreducible crystallographic root systems, the corresponding Weyl groups, and the simple Lie algebras are classified by the Dynkin diagrams of Fig. 7. Those having the roots of the same length (simply laced) are A_k, D_k, E_6, E_7, E_8. In fact, this was the reason why Arnold had chosen the notation for the simple singularities, see Sect. 3.2.1.

The root systems A_k and D_k correspond to the Lie algebras $sl(k+1)$ and $so(2k)$, respectively. The Weyl group A_k acts by permutations of coordinates in $\Pi_k \subset \mathbb{R}^{k+1}$, see the end of Sect. 3.2.2. The Weyl group D_k acts on $\mathbb{R}^k$ by permuting the coordinates and changing an even number of their signs. The root systems E_k, $k = 6, 7, 8$, are exceptional and have no simple description. The root system E_6 consists of 72 roots, E_7 of 126 roots, and E_8 of 240 roots that span $\mathbb{R}^6$, $\mathbb{R}^7$, and $\mathbb{R}^8$, resp.[13] These root systems define the corresponding simple Lie algebras and the Weyl groups, [29].

3.2.4 From Simple Singularities to Reflection Groups

Let $f \in \mathfrak{m}_n$ and $n \equiv 3 \pmod 4$. The intersection index in the integer homology of a non-singular level V_* defines a symmetric bilinear form, see Sect. 3.1.9.

Theorem 3.6 [11] *The simple singularities* (10) *are exactly the singularities possessing the non-degenerate sign-definite bilinear forms.*

The bilinear forms are negative definite. Therefore if a simple singularity f has multiplicity μ, then the *minus* intersection index defines a canonical Euclidean scalar product $(\cdot,\cdot) = -(\cdot \circ \cdot)$ on the homology $H_{n-1}(V_*; \mathbb{R}) \cong \mathbb{R}^\mu$ (with coefficients in $\mathbb{R}$!). The real monodromy group acting on $H_{n-1}(V_*; \mathbb{R})$ is generated by the Picard–Lefschetz operators (8) which appear to be the reflections in the hyperplanes orthogonal to the corresponding vanishing cycles. Clearly monodromy does not change the

[13] At http://en.wikipedia.org/wiki/Root_system#E8.2C_E7.2C_E6 the roots systems are nicely pictured.

intersection index. Thus the real monodromy group becomes a finite reflection group keeping the integer lattice $H_{n-1}(V_*) \subset H_{n-1}(V_*;\mathbb{R})$, i.e., it is a crystallographic reflection group.

Theorem 3.7 [11] *For each of the simple singularities* (10), *the basis of vanishing cycles in the homology of a non-singular level coincides with the set of simple roots of the root system of the same name*, *and the monodromy group is the Weyl group of this root system.*

By (7), all vanishing cycles have the same length, $\sqrt{2}$, hence for a simple singularity the root system is one of A_k, D_k, E_6, E_7, E_8. All the possibilities are realized, as the comparison of the lists shows.

Moreover for any simple singularity, the complex homology $H_{n-1}(V_*;\mathbb{C}) = H_{n-1}(V_*;\mathbb{R}) \otimes_{\mathbb{C}} \mathbb{C} \cong \mathbb{C}^\mu$ is the complexification of the Euclidean $H_{n-1}(V_*;\mathbb{R}) \cong \mathbb{R}^\mu$, together with the roots, the mirrors and the monodromy group action, exactly as in Arnold's example of Sect. 3.2.2. The orbit space is biholomorphically diffeomorphic to $\mathbb{C}^\mu$, and the image of mirrors under the Vieta map is the variety of irregular orbits (IO variety).

Theorem 3.8 [11] *For each of the simple singularities* (10), *the pair* $\{$*the base of a versal deformation*, *the discriminant*$\}$ *is diffeomorphic to the pair* $\{$*the orbit space*, *the IO variety*$\}$ *of the complexified action of the Weyl group of the same name.*

3.2.5 Hunting for Other Reflection Groups

In Arnold's paper [13] of 1978, the Weyl groups whose Dynkin diagrams have a double edge, B_k, C_k, F_4 of Fig. 7, came up in connection with the simple boundary singularities, in a quite similar to the $A - D - E$ case way.

Arnold's *"complexification of a manifold with boundary"*, that is, the double covering ramified along the boundary, provides a direct relation between the boundary singularities and the $\mathbb{Z}_2$-symmetric singularities. The only lacking Weyl group G_2 appears in connection with the only simple $\mathbb{Z}_3$-symmetric singularity.

In 1979, Arnold published a paper [14] where the relation between the simple singularities and the Weyl groups was made more deep and explicit. In particular, the local algebra Q_f of a simple singularity f of multiplicity μ was identified with the cotangent space at the origin, $T^*_O B$, to the orbit space $B = \mathbb{C}^\mu / W \cong \mathbb{C}^\mu$ of the complexified action of the corresponding reflection group W.

The Weyl groups are the irreducible finite reflection groups keeping a lattice. A wider class is provided by the *Coxeter groups*, i.e., the finite groups generated by Euclidean reflections. The additional, *non-crystallographic*, irreducible Coxeter groups are described by the Dynkin diagrams H_3, H_4 and $I_2(p)$, $p \geq 5$, see Fig. 7. The groups $I_2(p)$ are the symmetry groups of p-gons[14] in $\mathbb{R}^2$, H_3 is the symmetry group of the icosahedron in $\mathbb{R}^3$, and H_4 is the symmetry group of the 'hypericosahedron', a regular polytope in $\mathbb{R}^4$ with 120 vertices and 600 faces.

[14] The group $I_2(5)$ is denoted also by H_2.

The question arose: What singularities correspond to these Coxeter groups? Arnold wrote about that fascinating period, [16, Introduction]:

"The search of other reflection groups ($H_2, H_3, H_4; I_2(p)$) started immediately. During the fall 1982 the joint efforts of O.V. Ljashko, A.B. Givental, O.P. Shcherbak and the author [V. Arnold] led to the discovery of the icosahedron symmetry group H_3; it controls the singularities of the ray system and the fronts in the variational problem of fastest bypassing of a plane obstacle bounded by a generic curve with an inflection point… O.P. Shcherbak has found in 1984 the most complicated 'hypericosahedron' H_4, related to a singularity in the obstacle problem in 3-space". See [51] for details.

4 Lagrangian Singularities

In order to understand the unexpected cancelation of many terms in dull and long computations, the strange similarity of bifurcation diagrams in apparently unrelated problems and the mysterious appearance of the regular polyhedra in problems of applied mathematics, one has to replace the straightforward computations in differential geometry by the simple and general approach of symplectic and contact geometry.

V.I. Arnold, Singularities of caustics and wave fronts

Now, after a brief tour of singularity theory, let us return to Lagrangian submanifolds of a symplectic manifold (M^{2n}, ω) introduced in Sect. 2.

A fibration $M^{2n} \twoheadrightarrow B^n$ is called *Lagrangian* if it has Lagrangian fibers. For fixed n, all Lagrangian fibrations are locally equivalent to the standard example

$$T^*\mathbb{R}^n \cong \mathbb{R}^{2n} \twoheadrightarrow \mathbb{R}^n, \quad (p, q) \mapsto q,$$

where q is the coordinate on $\mathbb{R}^n$, and $T^*\mathbb{R}^n$ is the cotangent bundle with the Darboux coordinates (p, q), see (2).

Let $L \hookrightarrow M^{2n}$ be an immersed Lagrangian submanifold in the space of Lagrangian fibration $M^{2n} \twoheadrightarrow B^n$. The projection of L to B is called a *Lagrangian map*. The critical values of a Lagrangian map form its *caustic*.

A *Lagrangian equivalence* between two Lagrangian maps $L_1 \hookrightarrow M_1 \twoheadrightarrow B_1$ and $L_2 \hookrightarrow M_2 \twoheadrightarrow B_2$ is a symplectomorphism $M_1 \to M_2$ taking the fibers of $M_1 \twoheadrightarrow B_1$ to the fibers of $M_2 \twoheadrightarrow B_2$ and taking L_1 to L_2. A *Lagrangian singularity* is a germ of a Lagrangian map considered up to a Lagrangian equivalence. The caustic is determined by a Lagrangian singularity, up to a diffeomorphism of the base B^n.

The classification of Lagrangian singularities was reduced in [11] to that of families of functions. The key role here is played by the notion of a *generating function*. Namely, any Lagrangian submanifold $L \hookrightarrow T^*\mathbb{R}^n \twoheadrightarrow \mathbb{R}^n$ is given locally by a generating function $F(x, q)$ in the following way:

$$L = \{(p, q) | \exists x : \partial F/\partial x = 0, p = \partial F/\partial q\}. \tag{11}$$

Note that $F(x, q)$ is defined up to an additive constant, and therefore it can be treated as a deformation of $F(x, 0)$ in $\mathfrak{m}_n$, cf. the end of Sect. 3.1.5.

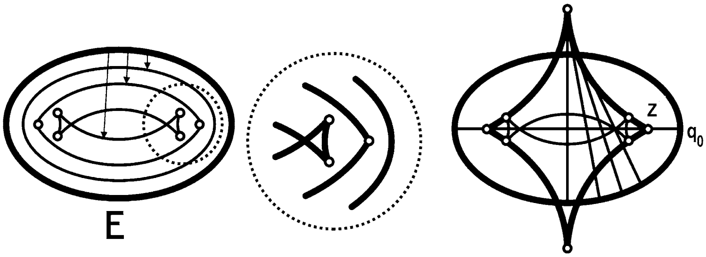


Fig. 9 Fronts and the caustic of an ellipse

Theorem 4.1 ([11, 12]) *Generating functions of generic Lagrangian singularities are truncated versal deformations of isolated singularities. In particular, any generic caustic at any point is diffeomorphic to the function bifurcation set of an isolated singularity.*

In particular, for $n \leq 5$ all Lagrangian singularities correspond to the truncated versal deformations of simple singularities. For $n = 2$, the caustic is a plane curve, and the only stable singularities are self-intersections and cusps.

Arnold's Second Example In yet another favorite example of Arnold, Lagrangian singularities appear in a wave propagation problem ([18, Introduction]).

Consider a velocity 1 wave propagation inside an ellipse E in the Euclidean plane $\mathbb{R}^2$. By definition, the *wave front* Φ_s is formed by the points $q + s\nu(q)$ where $q \in E$ and $\nu(q)$ is the inward normal to E at q. For small s the wave front is simply an equidistant of the ellipse. It is clearly smooth. When s increases, the front acquires self-intersection points and cusps, see Fig. 9. Later on, propagating beyond the ellipse, the wave front becomes smooth again. The cusps of the propagating front fill a curve called the *caustic*.[15] In the example the caustic is an astroid and has 4 cusps.

The "distance function" s is a multivalued function on the plane $\mathbb{R}^2$ whose level sets are the wave fronts. Consider its graph in the 3-dimensional space $\mathbb{R}^2_q \times \mathbb{R}_s$. It is obtained by lifting each wave front Φ_s to height s. Comparing Figs. 3 and 9 we observe that in a neighborhood of the focus of the ellipse, the graph is diffeomorphic to the swallowtail surface near the vertex, that is to the discriminant of A_3! This singularity is stable: after a slight perturbation of the ellipse, the graph of the distance function will have the same singularity at a nearby point. The caustic near the focal point is diffeomorphic to the bifurcation set of A_2 singularity, cf. Fig. 2.

We conclude that the surface formed by polynomials with a real multiple root in the 3-dimensional space of polynomials $\{x^4 + \lambda_1 x^2 + \lambda_2 x + \lambda_3\}$, near the point corresponding to x^4, governs the wave front propagation: when the wave front passes through the focus it changes the shape (bifurcates) in the same way as the sections of the swallowtail surface by planes $\lambda_1 = const$ change when λ_1 passes through 0 (Fig. 3).

[15] Geometrically, the caustic is the *envelope of the normal lines* to the ellipse. The self-intersection points of a propagating front fill another curve called the *Maxwell stratum* which is also studied in singularity theory, [18, Chap. 2], [23, 24, vol. II, Chap. 2, Sect. 3].

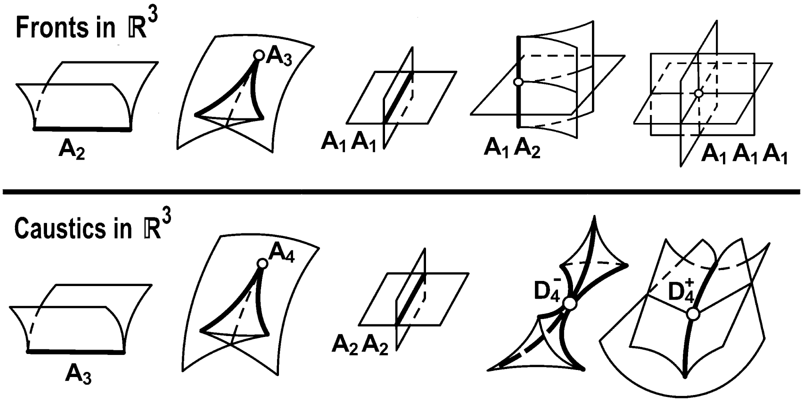


Fig. 10 Stable singularities of fronts and caustics in $\mathbb{R}^3$

In fact, the generic singularities of both fronts and caustics in $\mathbb{R}^3$ all are diffeomorphic to irregular orbit varieties of the Euclidean reflection groups,[16] including possible transversal intersections of these surfaces, as is shown on Fig. 10.

Now let us describe the wave propagation process above in symplectic terms. Again, at each point $q \in E$ of the ellipse consider its inward unit normal $\nu(q)$. Define a co-vector $\hat{\nu}(q) \in T_q^*\mathbb{R}^2$ by $\langle \hat{\nu}(q), \xi \rangle := (\nu(q), \xi)$ for all $\xi \in T_q\mathbb{R}^2$ (recall that $(\cdot, \cdot)$ is the scalar product). The curve $\widehat{E} := \{(\hat{\nu}(q), q), q \in E\}$ is a lift of the ellipse E to $T^*\mathbb{R}^2$. Let

$$g_s : T^*\mathbb{R}^2 \to T^*\mathbb{R}^2, \qquad (p, q) \to (p, q + sp)$$

be the Euclidean geodesic flow (recall that the wave propagation velocity is 1, so the distance s may be identified with the time). Under the action of the flow, the curve $\widehat{E}$ sweeps out a Lagrangian submanifold $L := \bigcup_s g_s(\widehat{E})$ in $T^*\mathbb{R}^2$.

The wave front Φ_s is given by $\pi(g_s(\widehat{E}))$, where $\pi : T^*\mathbb{R}^2 \to \mathbb{R}^2$ is the natural projection $(p, q) \mapsto q$. The caustic is the set of critical values of the projection $\pi|_L$.

Let us describe the Lagrangian singularity $L \hookrightarrow T^*\mathbb{R}^2 \twoheadrightarrow \mathbb{R}^2$ near the focus z of the ellipse E. Choose a parameterization $E(x)$, $x \in (-\epsilon, \epsilon)$, of an arch of the ellipse E so that $E(0) = q_0$ is the endpoint of the major semi-axis containing z, see Fig. 8. Then the distance function $F(x, q) := |q - E(x)|$ is a generating function of L, cf. (11). We conclude that the partial derivatives $\partial^k F/\partial x^k$ for $k = 1, 2, 3$ vanish at $(0, z)$, while the fourth derivative does not vanish.[17] Thus (cf. Theorem 4.1) the Lagrangian submanifold L near the preimage $(\pi|_L)^{-1}(z)$ of the focus z is equivalent to the one defined by the truncated versal deformation $G(x, \lambda) := x^4 + \lambda_1 x^2 + \lambda_2 x$ of the A_3-singularity. The caustic near the focus z is diffeomorphic to the function

[16] Here we deal with real analytic germs. Some complex singularities have non-equivalent real forms, e.g. the germs $D_4^{\pm}$ are given by $x_1^2 x_2 \pm x_2^3$, cf. a footnote in Sect. 3.1.2.

[17] The argument is as follows: the osculating circle centered at a generic point of the plane has tangency of order 2 with the ellipse; the centers of the osculating circles having tangency of order 3 form a line (the caustic); finally, some isolated points on this line are the centers of the osculating circles having tangency of order 4. In the example, there are four such points on the caustic, two of them are focuses of the ellipse.

bifurcation set of A_3. The latter is just the discriminant of the A_2-singularity, that is, a semi-cubical parabola, see the end of Sect. 3.1.5.

Further, given a point q near the focus and a value of the parameter x near 0 so that $\partial F/\partial x(x,q) = 0$, we have

$$q = E(x) + F(x,q) \cdot \nu((E(x)),$$

so (non-surprisingly!) the wave originating at $E(x)$ reaches q when $s = F(x,q)$. In other words, $q \in \Phi_s$, where $s = F(x,q)$.

Thus the set

$$W = \{(s,q) | \exists x \in (-\epsilon, \epsilon) : \partial F/\partial x(x,q) = 0,\ s = F(x,q)\} \subset \mathbb{R} \times \mathbb{R}^2$$

is the graph of the (multivalued) distance function. In view of the discussion above, this set is locally diffeomorphic to the discriminant of the A_3-singularity, that is, to a swallowtail surface.

Interestingly enough, generating functions of Lagrangian submanifolds in cotangent bundles serve as a powerful tool in global symplectic topology (see e.g. [33]).

Let X^n be a closed manifold and $L \subset T^*X$ be a closed Lagrangian submanifold. We say that $F : X^n \times \mathbb{R}^N \to \mathbb{R}$ is a generating function of L if (under certain additional regularity conditions on F)

$$L = \{(p,q) | \exists x : \partial F/\partial x(x,q) = 0,\ p = \partial F/\partial q(x,q)\}.$$

It turns out [30, 52] that if L is the image of the zero section X of T^*X under a Hamiltonian diffeomorphism, then L admits a generating function $F(x,q) = F_q(x)$ which is *quadratic at infinity in* x: outside a compact set, $F_q(x)$ is a non-degenerate quadratic form. Note that the critical points of $F_q(x)$ are in one-to-one correspondence with the intersection points $L \cap X$. Therefore, when L is transversal to X, an easy Morse-theoretical argument shows that the number of intersection points does not exceed the sum of Betti numbers of X.

The appearance of generating functions both in singularity theory and in symplectic topology manifests a fruitful interaction between these fields. In conclusion, let us mention two more examples of such an interaction. Symplectic features of the monodromy group of a singularity, highlighted in Arnold's paper [25], stimulated a discovery by Seidel [50] of a new class of symplectically knotted Lagrangian spheres in symplectic four-manifolds. Furthermore, there exists a profound interaction between singularity theory and mirror symmetry, for which we refer to Givental's article [36].

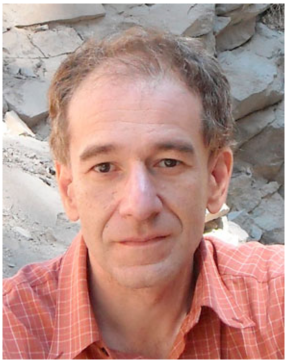
Acknowledgements We thank Hansjörg Geiges and Maxim Kazarian for valuable comments on the article as well as Natasha Artemeva, Miriam Hercberg and Julia Kreinin for their help in improving our English.

References

1. Arnold, V.I.: On Liouville's theorem concerning integrable problems of dynamics. Transl. Am. Math. Soc. **61**, 292–296 (1967) [Russian original: 1963]
2. Arnold, V.I.: Proof of a theorem by A.N. Kolmogorov on the persistence of quasi-periodic motions under small perturbations of the Hamiltonian. Russ. Math. Surv. **18**, 9–36 (1963)

3. Arnold, V.I.: Sur une propriété topologique des applications globalement canoniques de la mécanique classique. C. R. Acad. Sci. Paris **261**, 3719–3722 (1965)
4. Arnold, V.I.: Instability of dynamical systems with several degrees of freedom. Sov. Math. **5**, 581–585 (1964)
5. Arnold, V.I.: Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. Ann. Inst. Fourier **16**, 319–361 (1966)
6. Arnold, V.I.: On a characteristic class entering into conditions of quantization (Russian). Funct. Anal. Appl. **1**, 1–13 (1967)
7. Arnold, V.I.: A stability problem and ergodic properties of classical dynamical systems. In: Proc. Internat. Congr. Math. (Moscow), pp. 387–392 (1966)
8. Arnold, V.I.: The cohomology ring of the colored braid group. Math. Notes **5**, 138–140 (1969)
9. Arnold, V.I.: On the arrangement of the ovals of real plane curves, involutions of 4-dimensional smooth manifolds, and the arithmetic of integral quadratic forms. Funct. Anal. Appl. **5**, 169–176 (1971)
10. Arnold, V.I.: Modes and quasimodes. Funct. Anal. Appl. **6**, 94–101 (1972)
11. Arnold, V.I.: Normal forms for functions near degenerate critical points, the Weyl groups A_k, D_k, E_k and Lagrangian singularities. Funct. Anal. Appl. **6**, 235–272 (1972)
12. Arnold, V.I.: Mathematical Methods of Classical Mechanics. Graduate Texts in Mathematics, vol. 60. Springer, New York (1989). [Russian original: 1974]
13. Arnold, V.I.: Critical points of functions on a manifold with boundary, the simple Lie groups B_k, C_k, F_4 and singularities of evolutes. Russ. Math. Surv. **33**, 99–116 (1978)
14. Arnold, V.I.: Indices of singular points of 1-forms on a manifold with boundary, convolution of invariants of reflection groups, and singular projections of smooth surfaces. Russ. Math. Surv. **34**, 1–42 (1979)
15. Arnold, V.I.: First steps in symplectic topology. Russ. Math. Surv. **41**, 1–21 (1986)
16. Arnold, V.I.: Catastrophe Theory. In: Dynamical Systems V. Encyclopaedia of Mathematical Sciences, vol. 5. Springer, Berlin (1994) [Russian original: 1986]
17. Arnold, V.I.: Huygens & Barrow, Newton & Hooke. Birkhäuser, Basel (1990)
18. Arnold, V.I.: Singularities of caustics and wave fronts. In: Mathematic and Its Applications (Soviet Series), vol. 62. Kluwer Academic, Dordrecht (1990)
19. Arnold, V.I.: On the teaching of mathematics. Russ. Math. Surv. **53**, 229–236 (1998)
20. Arnold, V.I.: Yesterday and Long Ago. Springer, Berlin (2007). [Russian original: "Istorii davnie i nedavnie" 3rd edn, PHASIS, Moscow, 2006]
21. Demidovich, V.B.: Interview with V.I. Arnold. In: Mekhmatiane Vspominaut [Employees of the University Recollect]: 2 [Russian], pp. 25–58. Moscow State University, Moscow (2009)
22. Arnold, V.I., Avez, A.: Ergodic Problems of Classical Mechanics. Benjamin, New York (1968)
23. Arnold, V.I., Goryunov, V.V., Lyashko, O.V., Vasil'ev, V.A.: Singularity theory I: local and global theory. In: Dynamical Systems VI. Encyclopaedia of Mathematical Sciences, vol. 6. Springer, Berlin (1993). Singularity Theory II: Classification and Applications. In: Dynamical systems VIII. Encyclopaedia of Mathematical Sciences, vol. 8. Springer, Berlin (1993)
24. Arnold, V.I., Goryunov, V.V., Lyashko, O.V., Vasil'ev, V.A.: Singularity theory II: classification and applications. In: Dynamical Systems VIII. Encyclopaedia of Mathematical Sciences, vol. 8. Springer, Berlin (1993)
25. Arnold, V.I.: Some remarks on symplectic monodromy of Milnor fibrations. In: Hofer, H., Taubes, C., Weinstein, A., Zehnder, E. (eds.) The Floer Memorial Volume. Progr. Math., vol. 133, pp. 99–104. Birkhäuser, Boston (1995)
26. Arnold, V.I., Khesin, B.: Topological Methods in Hydrodynamics. Applied Mathematical Sciences, vol. 125. Springer, New York (1998)
27. Arnold, V.I., Kozlov, V.V., Neishtadt, A.I.: Mathematical aspects of classical and celestial mechanics. In: Dynamical Systems III. Encyclopaedia of Mathematical Sciences, vol. 3. Springer, Berlin (2006)
28. Buhovsky, L.: The Maslov class of Lagrangian tori and quantum products in Floer cohomology. J. Topol. Anal. **2**, 57–75 (2010)
29. Bourbaki, N.: Lie Groups and Lie Algebras. Elements of Mathematics. Springer, Berlin (2002). Chaps. IV–VI [French original: 1968]
30. Chekanov, Y.: Critical points of quasifunctions, and generating families of Legendrian manifolds. Funct. Anal. Appl. **30**, 118–128 (1996)
31. Chekanov, Y.: Lagrangian tori in a symplectic vector space and global symplectomorphisms. Math. Z. **223**, 547–559 (1996)

32. Conley, C.C., Zehnder, E.: The Birkhoff-Lewis fixed point theorem and a conjecture of V.I. Arnold. Invent. Math. **73**, 33–49 (1983)
33. Eliashberg, Y.: Symplectic topology in the nineties. Differ. Geom. Appl. **9**, 59–88 (1998)
34. Floer, A.: Morse theory for Lagrangian intersections. J. Differ. Geom. **28**, 513–547 (1988)
35. Floer, A.: Cuplength estimates on Lagrangian intersections. Commun. Pure Appl. Math. **42**, 335–356 (1989)
36. Givental, A.: Singularity theory and symplectic topology. In: The Arnoldfest, Toronto, ON, 1997. Fields Inst. Commun., vol. 24, pp. 201–207. Amer. Math. Soc., Providence (1999)
37. Givental, A.: http://math.berkeley.edu/~giventh/arnold_en.pdf
38. Gusein-Zade, S., Varchenko, A.: Vladimir Arnold. In: European Mathematical Society Newsletter No. 78, pp. 28–29 (December 2010)
39. Hofer, H.: Lagrangian embeddings and critical point theory. Ann. Inst. H. Poincaré, Anal. Non Lineaire **2**, 407–462 (1985)
40. Hofer, H., Zehnder, E.: Symplectic Invariants and Hamiltonian Dynamics. Birkhäuser, Basel (1994)
41. Kaloshin, V., Levi, M.: An example of Arnold diffusion for near-integrable Hamiltonians. Bull. Am. Math. Soc. **45**, 409–427 (2008)
42. Kolmogorov, A.N.: On the persistence of conditionally periodic motions under a small change of the Hamilton function. Dokl. Akad. Nauk SSSR **98**, 527–530 (1954). (in Russian). English translation: A.N. Kolmogorov. In: Casati, G., Ford J. (eds.) Stochastic Behavior in Classical and Quantum Hamiltonian Systems. Lecture Notes in Physics, vol. 93, pp. 51–56. Springer, Berlin (1979)
43. Laudenbach, F., Sikorav, J.-C.: Persistance d'intersection avec la section nulle au cours d'une isotopie hamiltonienne dans un fibre cotangent. Invent. Math. **82**, 349–357 (1985)
44. Maslov, V.P.: Théorie des perturbations et méthodes asymptotiques. Dunod, Paris (1972). [Russian original: 1965]
45. McDuff, D., Salamon, D.: Introduction to Symplectic Topology. Oxford Mathematical Monographs. Clarendon Press, New York (1998)
46. McDuff, D., Salamon, D.: J-Holomorphic Curves and Symplectic Topology. Amer. Math. Soc., Providence (2004)
47. Moser, J.: On invariant curves of area preserving mappings of an annulus. Nachr. Akad. Wiss. Gött., Math. Phys. Kl. 1–20 (1962)
48. Polterovich, L.: https://sites.google.com/site/polterov/home/remembering-vladimir-arnold/
49. Schwarz, M.: Morse Homology. Progress in Mathematics, vol. 111. Birkhäuser, Basel (1993)
50. Seidel, P.: Lagrangian two-spheres can be symplectically knotted. J. Differ. Geom. **52**(1), 145–171 (1999)
51. Shcherbak, O.: Wave fronts and reflection groups. Russ. Math. Surv. **43**, 149–194 (1988)
52. Sikorav, J.-C.: Problèmes d'intersections et de points fixes en géométrie hamiltonienne. Comment. Math. Helv. **62**, 62–73 (1987)
53. Weinstein, A.: Lectures on symplectic manifolds. In: Expository Lectures from the CBMS Regional Conference Held at the University of North Carolina, 8–12 March, 1976. Regional Conference Series in Mathematics, vol. 29. Am. Math. Soc., Providence (1977)
54. Yakovenko, S.: http://yakovenko.wordpress.com/2010/06/12/vladimir-igorevich-arnold-sad-birthday/
55. Zehnder, E.: Generalized implicit function theorems with applications to some small divisor problems I, II. Commun. Pure Appl. Math. **28**, 91–140 (1975); **29**, 49–111 (1976)



Leonid Polterovich participated in Sinai's and Arnold's seminars at Moscow State Lomonosov University until his emigration to Israel in 1990. He completed his PhD at Tel Aviv University in 1990. Currently Polterovich holds the Gordon Chair in Symplectic Topology and Hamiltonian Dynamics at Tel Aviv University and a visiting professorship at University of Chicago.

Inna Scherbak was a graduate and post-graduate student of Vladimir Arnold at Moscow State Lomonosov University. She got her PhD in 1986, for the thesis "Lagrange and Legendrian boundary singularities" written under Arnold's supervision. She is working in singularity theory and algebraic geometry.

Jahresber Dtsch Math-Ver (2011) 113:221–224
DOI 10.1365/s13291-011-0030-y

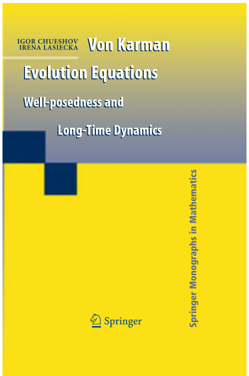
BUCHBESPRECHUNG

Igor Chueshov and Irena Lasiecka: “Von Karman Evolution Equations”

Springer-Verlag, 2010, 778 pp.

Albert J. Milani

Published online: 21 October 2011



1. This book is devoted to a mathematical presentation of a simplified version of the so-called “von Karman” equations, which model the dynamics of the vertical oscillations of an elastic two-dimensional plate, due to various types of forces. The general form of these equations is as follows. Let Ω be a bounded domain of $\mathbb{R}^2$ and denote its points by $x = (x_1, x_2)$. Given two functions $v,\ w : \Omega \to \mathbb{R}$, define a third function formally by

$$[v, w] := v_{11}w_{22} + v_{22}w_{11} - 2v_{12}w_{12} \tag{1}$$

where $v_{ij} := \partial_i \partial_j v$, and note that, if $w = v$,

$$[v, v] = 2\det \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}; \tag{2}$$

that is, $[v, v]$ is twice the determinant of the Hessian matrix of v. In addition, denote by $\Delta := \partial_1^2 + \partial_2^2$ the Laplace operator in Ω. Given a non-negative parameter $\alpha \in \mathbb{R}_{\geq 0}$, two functions $F_0 = F_0(t, x)$ and $p = p(t, x)$ defined on $Q := \mathbb{R}_{\geq 0} \times \Omega$, and a first order differential operator L with smooth coefficients in Q, one tries to determine a function $u = u(t, x)$ satisfying the system of partial differential equations (PDEs)

$$u_{tt} - \alpha\Delta u_{tt} + \Delta^2 u - [u, v + F_0] + Lu = p, \tag{3}$$
$$\Delta^2 v + [u, u] = 0. \tag{4}$$

A.J. Milani (✉)
University of Wisconsin, Milwaukee, USA
e-mail: ajmilani@uwm.edu

In this system, the unknown u represents the vertical displacement of the plate; v represents the so-called "Airy stress function", which is related to the internal elastic forces acting on the plate, and depends on the deformation u of the plate, as described by (4). The four terms $\alpha\Delta u_{tt}$, $[u, F_0]$, Lu and p of (3) are related to different types of forces acting on the plate; specifically to rotational inertial forces, internal forces, non-conservative feed-back forces, and transverse forces, respectively. Note that the model (3) + (4) does not take into consideration the accelerations due to horizontal displacement of the plate. Equation (4) is supplemented by the boundary conditions

$$v|_{\partial\Omega} = 0, \qquad \nabla v|_{\partial\Omega} = 0, \tag{5}$$

while u satisfies one of three different types of boundary conditions (BC); namely, either the CLAMPED BC

$$u|_{\partial\Omega} = 0, \qquad \nabla u|_{\partial\Omega} = 0 \tag{6}$$

(as in (5)), which require that the plate be fixed and remain horizontal at the boundary; or the HINGED BC

$$u|_{\partial\Omega} = 0, \qquad \Delta u|_{\partial\Omega} = 0, \tag{7}$$

which require that the plate be fixed, but allowed to pivot at the boundary (the term Δu in (7) represents the so-called "bending moment" of the plate); or the FREE BC

$$\begin{aligned} (\Delta u + (1-\mu)B_1 u)|_{\partial\Omega} &= 0, \\ (n \cdot \nabla(\Delta u - \alpha u_{tt}) + (1-\mu)B_2 u - \nu u)|_{\partial\Omega} &= 0, \end{aligned} \tag{8}$$

where $\mu \in [0, 1]$ and $\nu \geq 0$ are constants with a specific physical meaning; B_1 and B_2 are second-order boundary differential operators, naturally arising from integration by parts, and n denotes the outward normal to $\partial\Omega$. Numerous variations of these BC are also considered, including nonlinear ones. The BC may also be mixed, in the sense that u should satisfy different BC on different parts of $\partial\Omega$. Finally, u is subject to the initial conditions

$$u(0, x) = u_0(x), \qquad u_t(0, x) = u_1(x), \tag{9}$$

where u_0 and u_1 are given functions on Ω.

2. The resulting initial-boundary value problems (IBVPs) are first reformulated in abstract form in a suitable frame of Banach spaces, and solved by means of either non-linear semigroup theory or Galerkin approximation schemes. Depending on the assumptions on the data, the solutions obtained are strong, weak, or generalized; however, they are global in time, at least in the sense that they are defined on finite, but arbitrary time intervals $[0, T]$. More precisely, strong (respectively, weak) solutions of (3) + (4), are pairs (u, v) such that

$$(u, v) \in C^0([0, T]; X_1 \times X_2) \cap C^1([0, T]; Y_1 \times Y_2) \tag{10}$$

(respectively,

$$(u, v) \in W^{0,\infty}(0, T; X_1 \times X_2) \cap W^{1,\infty}(0, T; Y_1 \times Y_2)), \tag{11}$$

where, for $i = 1, 2$, X_i and Y_i are Banach spaces of functions on Ω, with $X_i \hookrightarrow Y_i$; generalized solutions are generally defined as limits, in a suitable topology, of sequences of strong solutions. An analogous representation holds for solutions of systems of more than two unknowns. In general, the spaces X_i and Y_i are closed subspaces of suitable Sobolev spaces on Ω, whose degree of regularity is also part of the distinction between weak and strong solutions. The challenge is, of course, to determine the spaces X_i and Y_i in such a way as to conveniently describe the various types of boundary conditions; for example, for strong solutions of the IBVP with the hinged BC (7) for u, and the clamped BC (5) for v, it is natural to choose

$$\begin{cases} X_1 = \{u \in H^3(\Omega) \cap H_0^1(\Omega) \mid \Delta u \in H_0^1(\Omega)\}, \\ Y_1 = H^2(\Omega) \cap H_0^1(\Omega), \end{cases} \tag{12}$$

$$X_2 = H_0^2(\Omega), \qquad Y_2 = H_0^1(\Omega). \tag{13}$$

In the *a priori* estimates that are required to implement these methods, a fundamental role is played by the properties of the bracket $[v, w]$; more precisely, to estimate the products appearing in (1), the usual Sobolev product estimates need to be supplemented by more refined estimates in appropriate Besov-Lizorkin-Hardy spaces. This solution theory is presented in the first half of the book, together with a number of extensions, such as, for example, thermo-elastic plates, when the deformation of the plate is caused also by heat (in the corresponding model, (3) is coupled with a linear heat equation describing the evolution of the temperature of the plate), or plates immersed in a gas, in which the pressure p in (3) depends in various way on the velocity potential of the flowing gas, which in turn also depends on the displacement u and its first order derivatives.

3. Since the solutions of the IBVPs are defined for all $t \geq 0$, each IBVP defines a semiflow $S = (S(t))_{t \geq 0}$ of operators (in general, nonlinear) $S(t) : Z \to Z$, where Z is a suitable Banach space, called the "phase space" of the semiflow. This means that: (i) $S(0) = I$; (ii) $S(t_1 + t_2) = S(t_1)S(t_2) = S(t_2)S(t_1)$; (iii) For each $z_0 \in Z$, the map $t \mapsto S(t)z_0$ is continuous from $[0, +\infty[$ into Z; (iv) For each $t \geq 0$, the map $S(t)$ is continuous from Z into itself. The set $(S(t)z_0)_{t \geq 0}$ is called the forward trajectory starting at z_0. For example, system (3) + (4) can be rewritten as a single equation, replacing into (3) the functional dependence $v = \Phi(u)$ defined by (4). In this case, $Z = X_1 \times Y_1$, and if u is the strong solution of this equation, with initial values u_0 and u_1 as in (9), and the hinged BC (7), the corresponding semiflow is defined by

$$S(t)(u_0, u_1) = (u(t), u_t(t)). \tag{14}$$

Indeed, condition (i) is trivial; condition (ii) holds because the equations are autonomous; condition (iii) holds, because the solution is differentiable in t, and condition (iv) is a consequence of the continuity of the solution with respect to its initial values. Of course, these conditions (in particular, the last one), have to be checked in each example. The second part of the book is devoted to the asymptotic properties, as $t \to +\infty$, of the solutions of each IBVP, which are studied in terms of the long-time behavior of the trajectories, considered as subsets of the phase space Z. In the context of semiflow theory, this is usually achieved by showing the existence of some subset of Z which attracts the orbits (in a suitable sense). The three most important attracting sets, which are also the sets considered in this book, are the global attractors, the exponential attractors, and the inertial manifolds. Roughly speaking, the global attractor is a compact set $\mathcal{A} \subset Z$, invariant under S (that is, $S(t)\mathcal{A} = \mathcal{A}$ for all $t \geq 0$), which has finite fractal dimension and attracts all orbits; an exponential attractor is a compact set $\mathcal{E}$, positively invariant under S (that is, $S(t)\mathcal{E} \subseteq \mathcal{E}$ for all $t \geq 0$), which has finite fractal dimension and attracts all bounded sets with an exponential rate; an inertial manifold is a finite-dimensional Lipschitz submanifold of Z, which is positively invariant and attracts all trajectories exponentially. The fact that an attractor has a finite dimension allows for "reducing" the study of the asymptotic behavior of an infinite dimensional dynamical system, such as those generated by PDEs, to finite-dimensional ones, that is, essentially, to a system of ODEs (this is, in essence, the spirit of Mañé's theorem). A necessary condition for the existence of all these attracting sets is the existence of a bounded, positively invariant attracting set; this property is usually known as "dissipativity" of the semiflow, and generally follows from some sort of damping term, present in the PDEs, or in the BC. The first two chapters of the second part of the book present the main properties of semiflows, and a number of sufficient conditions on a dissipative semiflow, that guarantee the existence of some of these attracting sets. In the remaining chapters, it is shown that the dissipative semiflows generated by the dynamical systems corresponding to the IBVPs studied in the first part of the book do satisfy some or all of these conditions, and therefore admit one or all of the corresponding attracting sets. Included are many results on the estimate of the dimension of the global and the exponential attractors.

4. In conclusion, this book is a comprehensive presentation of virtually all that is presently known on the von Karman evolution equations in two space dimensions, together with the necessary theoretical background on semiflows. As such, it requires a somewhat demanding effort from the reader, whose tenacity, however, will be rewarded by the wealth of information that this book supplies and provides.

Jahresber Dtsch Math-Ver (2011) 113:225–227
DOI 10.1365/s13291-011-0028-5

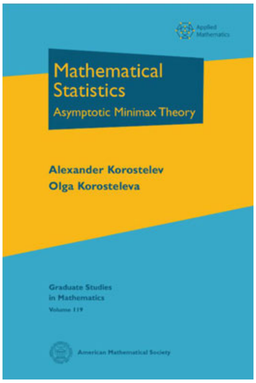
BUCHBESPRECHUNG

Alexander Korostelev, Olga Korosteleva: “Mathematical Statistics: Asymptotic Minimax Theory”

AMS, 2011, 246 pp.

Winfried Stute

Online publiziert: 18. Oktober 2011



Mathematische Resultate werden gewöhnlich mit einer Beschreibung der Rahmenbedingungen in Form von „Gegeben sei …“ eröffnet. Man möge sich einmal vorstellen, was geschähe, wenn diese Art von Selbstbedienung nicht greift und Studierenden bzw. Anwendern relevante Informationen vorenthalten würden. Eine absurde Vorstellung? Nicht ganz, sondern im Gegenteil Alltag des Statistikers. Statt präziser Vorgaben liegt die Information in der Regel lediglich in Form von diskreten Daten vor, und das Ziel könnte darin bestehen, daraus möglichst fehlerfrei interessierende Funktionen zurückzugewinnen. Die vorliegende Monographie beschäftigt sich in 15 Kapiteln eingehend mit der Fehleranalyse von statistischen Schätzverfahren in unterschiedlichen Szenarien. Testprobleme werden (in Kapitel 16) nur am Rande behandelt. Teil 1 diskutiert in 7 Kapiteln Schätzverfahren in parametrischen Modellen, Teil 2 konzentriert sich auf Fragestellungen im Rahmen der Nichtparametrischen Regression (Kapitel 8–12), während in Teil 3 spezielle Probleme bei der Nichtparametrischen Modellierung angesprochen werden.

Parametrische Modelle zeichnen sich dadurch aus, dass zu schätzende Funktionen bis auf einen unbekannten endlich-dimensionalen Parameter vollständig festgelegt sind. Die Autoren beschränken sich grundsätzlich auf den Fall eindimensionaler Parameter, was man didaktisch begründen kann. Letztendlich ist man damit jedoch nicht in der Lage, einfachste Normalverteilungsmodelle adäquat zu behandeln. Ausgangspunkt der Untersuchungen ist die Cramér-Rao-Ungleichung, die eine untere Schranke

W. Stute (✉)
Gießen, Deutschland
e-mail: Winfried.Stute@math.uni-giessen.de

für den Mean Square Error (MSE) eines Schätzers (im regulären Fall) darstellt. Wird diese Schranke angenommen, kann man sich sicher sein, dass man damit das Optimum erreicht hat. In vielen Fällen hängt diese Schranke vom Parameter ab, so dass die Frage auftaucht, ob anstelle eines punktuellen Vergleichs nicht integrierte Fehler (Bayes Risiko) oder maximale Fehler zu betrachten seien. Will man den maximal zu erwartenden Schaden minimieren, führt dies zu Fragestellungen, die der Monographie ihren Namen geben (Minimaxity). Kapitel 2 behandelt den Zusammenhang zwischen Bayes und Minimax Schätzern, und Kapitel 3 diskutiert die asymptotische Variante. Die im Detail diskutierten Beispiele bzw. Modelle sind klassisch und entstammen speziellen Exponentialfamilien. Kapitel 4–7 behandeln eher ausgewählte Fragestellungen. So gehen die Autoren in Kapitel 4 an einigen Beispielen (wie z. B. den Gleichverteilungen) auf irreguläre Modelle ein, während Kapitel 5 sich mit dem Aufdecken von sogenannten Change Points beschäftigt. Dies sind unbekannte Zeitpunkte, zu denen Verteilungsfunktionen, Dichten oder Parameter sich verändern. In der Regel werden wieder einfache parametrische Modelle (mit eindimensionalen Parametern) betrachtet. Kapitel 6 spricht im Rahmen der Minimax Theorie sequentielle Schätzungen an, wo also der Stichprobenumfang zu Beginn nicht geplant ist und z. B. das Ziel darin besteht, ein zu frühes Stoppen (d. h. Anzeigen eines Wechsels) in Form einer False Alarm Probability zu kontrollieren. Kapitel 7 diskutiert einige klassische Aussagen zum Kleinst-Quadrate-Schätzer im Linearen Regressionsmodell. Insgesamt hat man den Eindruck, dass die einzelnen Themen nur einführend in einfachen analytisch handhabbaren Situationen behandelt werden.

Der weitaus größte Teil der Monographie ist den Teilen 2 und 3 gewidmet. Wurden in Teil 1 ausschließlich parametrische Modelle diskutiert, so sind es nun nichtparametrische. Schwerpunkt dabei bildet die nichtparametrische Regressionsschätzung, also die Rekonstruktion einer unbekannten Funktion aus einem Scatterplot. Mathematisch gesehen handelt es sich dabei um ein schlecht gestelltes Problem. Minimax Theorie in diesem Zusammenhang bedeutet, Schätzer zu finden, die hinsichtlich eines nichtparametrischen Modells, welches sich in der Regel nur durch Annahmen wie Lipschitz-Stetigkeit, Differenzierbarkeit etc. auszeichnet, den größten anzunehmenden Fehler zu minimieren. Die Untersuchungen haben das Ziel, optimale Konvergenzraten zu bestimmen. Durch Fokussierung auf den schlimmsten anzunehmenden Fehler geraten andere Größen wie optimale Konstanten, die von unbekannten Größen abhängen und die Komplexität eines schlecht gestellten Problems mit ausmachen, in den Hintergrund. Dies mag man bedauern. Ansonsten werden einige klassische Glättungsverfahren diskutiert, wie Kernschätzer, Splines oder Orthogonalreihen-Schätzer. Darüber hinaus werden sowohl Resultate für festes Design als auch für zufälliges Design präsentiert. Als Fehlergrößen kommen lokale Abweichungen sowie globale (z. B. L^2-Abstände) zur Sprache.

Teil 3 schließlich spricht kurz einige semiparametrische Modelle (wie die additive Regression oder das Single-Index-Modell) an.

Insgesamt gesehen bleibt nach Lesen dieser Monographie ein zwiespältiges Gefühl zurück. Durchgängig überwiegt der Eindruck, nur aus einer speziellen Perspektive und nicht einmal halbwegs vollständig über relevante Fragen informiert worden zu sein. Viele Kapitel sind kurz und tippen die jeweiligen Themen nur an, häufig in einem Szenario, was auf den kritischen Betrachter sehr speziell wirkt. Zum Schluss bleibt die Frage zu beantworten, ob es sich bei dieser Monographie um ein Werk handelt, das beiden Komponenten im Titel (Mathematik und Statistik) einigermaßen gerecht wird. Die Antwort ist eindeutig, wenn man weiß, dass es die Autoren nicht einmal schaffen, ihre mathematischen Ergebnisse anhand einer Simulation, geschweige denn an einem realen Datensatz, zu illustrieren.

Jahresber Dtsch Math-Ver (2011) 113:229–233
DOI 10.1365/s13291-011-0029-4

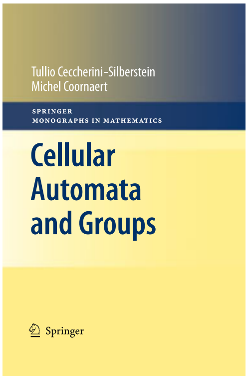
BUCHBESPRECHUNG

Tullio G. Ceccherini-Silberstein, Michel Coornaert: "Cellular Automata and Groups"

Springer-Verlag, 2010

Laurent Bartholdi

Published online: 21 October 2011



This in-depth book describes connections between apparently unrelated topics: cellular automata, group theory, and ring theory. I will focus on some striking results therein; remarkably, three of the four are all related to work of John von Neumann.

1 Cellular Automata

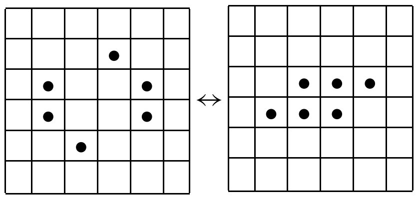
In his attempts at understanding life and intelligence, John von Neumann devised *cellular automata* as complex organisms made of simple components. In biology, they would be corals (class *anthozoa*); in mathematics, they consist of a graph (vertices called *cells*, and *edges*), a finite set of *states* in which each cell may be, and an *evolution rule* specifying in which state a cell should become, dependent on its current state and the states of its neighbours in the graph. The graph is assumed to have enough regularity that all neighbourhoods look the same; moreover, the evolution rule is the

L. Bartholdi (✉)
Göttingen, Deutschland
e-mail: laurent.bartholdi@gmail.com

same at each vertex. The metaphor holds because every vertex is thought to hold a microorganism, which may only interact with its immediate neighbours.

Von Neumann wanted the graph to be a three-dimensional lattice, so as to better describe the physical world. It was soon understood that a two-dimensional lattice could already give rise to extremely complex cellular automata.

One of the simplest, yet most fascinating examples is the *game of life*, due to John H. Conway. In this Manichean setting, each cell may be *alive* or *dead*. The graph is the usual square grid, with diagonals added, so that each cell has eight neighbours. A cell resuscitates if exactly three of its neighbours are alive, and survives if two or three of them are alive; otherwise, it dies from loneliness or overcrowding. For example,



Cellular automata may be used to implement complex computations using simple constituents. In fact, Conway's game of life is *Turing-complete*: any calculation that may be performed on a computer may be readily encoded into an initial configuration of the grid, in such a way that the result of the calculation may be reread from the grid at a later time.

This implies that most questions one may want to ask about the long-term behaviour of a cellular automaton do not have algorithmically-obtainable answers: "will all cells eventually die?"; "will the organism keep growing indefinitely?"; "will two given organisms eventually evolve to the same?"; etc.

2 Groups

The book under review advocates a more algebraic approach to cellular automata. Let G be a finitely generated group, and fix a generating set S for G. As underlying graph of the cellular automaton, consider the *Cayley graph* of G, that is, the graph with vertex set G, and an edge from g to gs for all $g \in G$ and all $s \in S$.

If A be the set of states of a cell in the automaton, the global state of the cellular automaton is described by a function $G \to A$. The evolution of the automaton is described by a map $A^{S \cup \{1\}} \to A$; by left-translation in G, it gives a map on functions $A^G \to A^G$. For example, the game of life of the previous section corresponds to $G = \mathbb{Z}^2$, generated by $S = \{(0, \pm 1), (\pm 1, 0), (\pm 1, \pm 1)\}$. So as to obtain the proper combination of global complexity and local simplicity, we wish G to be infinite, but S to be finite.

The universe of infinite, finitely generated groups is split in two classes: *amenable* and *non-amenable* groups. This notion, also due to von Neumann, is defined as follows: a group G is amenable if it admits a *mean*, that is, if a number $\mu(X)$ may be assigned to every subset X of G in such a way that $\mu(G) = 1$, $\mu(X \sqcup Y) =$

$\mu(X) + \mu(Y)$, and $\mu(gX) = \mu(X)$ for all $g \in G$. There are sundry equivalent definitions of amenability, and one of the most useful ones is the *fixed-point property*: every affine action of G on a non-empty convex compact set admits a fixed point.

Finite groups are amenable (set $\mu(X) = \#X/\#G$); so are Abelian groups. On the other hand, it is difficult to describe explicitly means on infinite groups. One may define a mean on $\mathbb{Z}$ by $\mu(X) = \lim_\omega \#(X \cap \{1, \ldots, n\})/n$, but it requires the existence of a non-principal ultrafilter ω so as to make every bounded sequence converge (the ordinary limit need not converge). The class of amenable groups is closed under taking subgroups, extensions, quotients, and limits.

Non-amenable groups, on the other hand, admit *paradoxical decompositions*: if G is not amenable, then there exists a partition $G = A_1 \sqcup \cdots \sqcup A_m \sqcup B_1 \sqcup \cdots \sqcup B_n$ such that, using left translations, $G = g_1 A_1 \sqcup \cdots \sqcup g_m A_m$ and $G = h_1 B_1 \sqcup \cdots \sqcup h_n B_n$. In other words, G can be cut into finitely many pieces in such a manner that the pieces can be rearranged using left translations, giving two copies of the original group. This is at the heart of the *Hausdorff-Banach-Tarski paradox*, that claims that a ball (say, made of solid gold) may be cut into finitely many pieces that, after rotation, may be fit into two copies of the original ball, thus doubling the amount of gold. Indeed, the free group F_2 is non-amenable, and acts by rotations on the ball, with essentially free orbits.

3 Gardens of Eden

Two properties of cellular automata have been singled out as being of particular interest. On the one hand, the evolution map of the automaton may fail to be injective, in which case there exist two distinct configurations of the automaton that differ in finitely many places, and that evolve in one step to the same configuration. These configurations, or rather their restrictions to where they differ, are called *mutually erasable patterns*.

On the other hand, the evolution map may fail to be surjective; in this case, there exists a pattern (restriction of a configuration to a finite subset of the group) that cannot appear through evolution. Such a pattern is called a *garden of Eden*, in reference to a paradisaical state of the universe that can never be reached again.

For example, it is clear that the game of life has mutually erasable patterns: a 3×3 dead grid and a 3×3 grid with a lone live cell in the middle have the same evolution. Existence of gardens of Eden is harder to ascertain.

It is nevertheless a theorem of Moore and Myhill that, if $G = \mathbb{Z}^2$, then a cellular automaton admits gardens of Eden if and only if it admits mutually erasable patterns.

The authors explain how this result is in fact related to amenability: if G is amenable, then a cellular automaton admits gardens of Eden if and only if it admits mutually erasable patterns. Conversely, if G is non-amenable, then there exist cellular automata which admit mutually erasable patterns but no gardens of Eden.

4 Surjunctive and Sofic Groups

Call a group G *surjunctive* if, for every cellular automaton on G, its evolution map is surjective as soon as it is injective. Obviously, this holds for finite G, since every injective self-map on a finite set is bijective. Injective maps have no mutually erasable patterns, so the results from the previous paragraph imply that amenable groups are surjunctive.

A group is residually finite if its elements can be distinguished in finite quotients; that is, if $g \neq h \in G$, then there exists a finite quotient $\pi : G \to Q$ such that $\pi(g) \neq \pi(h)$. For example, free groups are residually finite, while $\mathbb{Q}$ is not (it doesn't even have a non-trivial finite quotient). The authors show that residually finite groups are surjunctive, and give descriptions of that class of groups.

Given a class $\mathcal{C}$ of groups (say, finite groups), a group G is *locally embeddable* in $\mathcal{C}$ if, for every finite subset $K \subset G$, there exists a group $C \in \mathcal{C}$ and an injective map $\phi : K \to C$ satisfying $\phi(k_1k_2) = \phi(k_1)\phi(k_2)$ whenever $k_1, k_2, k_1k_2 \in K$. In other words, arbitrarily large finite parts of G can be approximated by groups in $\mathcal{C}$.

The authors consider next the following important relaxation of local embeddability. Let F be a finite set. Endow the symmetric group $\mathfrak{S}_F$ with the Hamming metric: $d(\alpha, \beta) = \frac{\#\{f \in F | \alpha(f) \neq \beta(f)\}}{\#F}$. A group G is *sofic* if for every finite subset $K \subset G$ and every $\epsilon > 0$ there exists a map $\phi : K \to \mathfrak{S}_F$ satisfying $d(\phi(k_1k_2), \phi(k_1)\phi(k_2)) \leq \epsilon$ whenever $k_1, k_2, k_1k_2 \in K$ and $d(\phi(k_1), \phi(k_2)) \geq 1 - \epsilon$ whenever $k_1 \neq k_2 \in K$. (The terminology comes from the Hebrew סופי, meaning 'finite'.) In other words, arbitrarily large finite parts of sofic groups can be approximated, with arbitrarily small error, by symmetric groups with their Hamming distance.

The class of sofic groups is closed under various operations (sums, limits, subgroups, ...). Groups locally embeddable in amenable groups are sofic. One of the major open questions in geometric group theory is whether *all* groups are sofic.

This question is intimately related to cellular automata, because sofic groups are surjunctive. The panorama is:

$$\begin{array}{ccccccc}
\text{finite} & \Longrightarrow & \text{res. finite} & \Longrightarrow & \text{loc. embed. finite} & & \\
\Downarrow & & \Downarrow & & \Downarrow & & \\
\text{amenable} & \Longrightarrow & \text{res. amenable} & \Longrightarrow & \text{loc. embed. amenable} & \Longrightarrow & \text{sofic} \\
 & & & & & & \Downarrow \\
 & & & & & & \text{surjunctive.}
\end{array}$$

5 Group Rings

Assume finally that the state set A of each cell in the cellular automaton is a finite-dimensional $\Bbbk$-vector space. Then the space of configurations A^G is also a $\Bbbk$-vector space; the evolution is required to be $\Bbbk$-linear. Assuming S is finite, this means the evolution map is given by convolution with a finitely-supported matrix-valued function on G, that is, a matrix over the group ring $\Bbbk G$.

In the last chapter, the authors explain how classical notions in ring theory translate to properties of cellular automata. Consider for example the following notion, also due to von Neumann: a ring R is *directly finite* if, whenever $a, b \in R$ satisfy $ab = 1$, we also have $ba = 1$. It is *stably finite* if all matrix rings over R are directly finite.

This notion is connected to linear cellular automata in the following manner: the group ring $\Bbbk G$ is stably finite if and only if, for every linear cellular automaton over G, the evolution map is surjective as soon as it is injective.

6 Conclusion

This remarkable book combines three, at first sight unrelated, concepts all due to von Neumann: cellular automata, amenability, and direct finiteness. The notions are explained in great detail, with numerous examples and historical remarks. The text is supplemented by ten appendices, recalling and detailing classical mathematical concepts (from abstract topology, group theory, functional analysis) that may equally serve as a reference as for "hands-on" learning.