INTRODUCTION TO SINGULARITY THEORY

Michail Zhitomirskii

Department of Math, Technion, 32000 Haifa, Israel

Lecture 1 Ideology, main notions and definitions

Lecture 2 Singularities of plane curves

Lecture 3 Constructing mini-versal deformations. Infinitesimal method.

These three lectures were given in Trieste, SISSA, on September 30, October 5, and October 6, 2010. Many thanks to Professor Andrei Agrachev for inviting me and for very useful for me math discussions. Many thanks to all listeners for questions, comments, and patience (the lectures were twice longer than scheduled). I am very thankful to Alesha Remizov who took all the job on publishing these notes.

Lecture 1. Ideology, terminology, main notions and definitions

I will start (section 1) with explanation of the most popular in singularity theory word "generic". In section 2 I will explain the first principle: which results are worth to obtain and which are not. In section 3 I define singularity classes in classification problems and their codimensions. In sections 4 and 5 I continue the main notions defining normal forms and modality. In section 6 I explain the second principle on the necessity to study families when attacking a non-generic singularity class. In section 7 I define versal deformations. The last section 8 is an example related to the Gauss method for solving systems of linear equations illustrating the significance of versal deformations. A good part of other illustrating examples also concern linear algebra and only few ones concern local analysis. A number of examples from local analysis will be given in Lectures 2 and 3. Unfortunately in three lectures I have no time to explain the Andronov-Hopf bifurcation in dynamical systems in terms of versal deformations - one of the most bright application of the ideology of singularity theory.

1. What does "generic" mean

I will start with the most commonly used word in singularity theory, the word "*generic*". To explain this word let me give few examples of sentences that have no sense:

1.1. Let A be a generic 2×2 matrix

1.2. Let $f(x)$ be a generic function.

And now the sentences (statements) with "generic" that have a precise sense:

2.1. A generic 2×2 matrix is non-singular

2.2. A generic 2×2 matrix has non-zero trace

2.3. If $f(x)$ is a generic C^∞ function and $f(x_0) = 0$ then $f'(x_0) \neq 0$.

From these examples you see that "generic" is not a characteristic of an individual object; *it is characteristic of a property meaning that this property holds for open and dense set of objects.* Statement 2.1 , respectively statement 2.2 is a short form to say that the set of non-singular 2×2 matrices, respectively the set of 2×2 matrices with non-zero trace, is open and dense in the space of all 2×2 matrices. Statement 2.3 is a short way to say that union of the set of C^∞ functions which do not vanish at any point with the set of C^∞ functions which vanish at some point or points and whose derivative does not vanish at these points is open and dense in the space of all C^∞ functions.

We see that "generic" requires topology. In statements 2.1 and 2.2 we deal with a finite dimensional vector space of objects and there is no problem with topology; these statements are obvious. Statement 2.3 is not that obvious and whether it is true or not depends on topology. If we want it to be true the C^0 topology is not enough; C^1 topology is enough if the argument x varies on a compact set; if it varies on the whole line $\mathbb{R}$ then the statement 2.3 is true only in a rather special topology called weak Whitney topology.

2. Principle 1: the order of obtaining results

Studying the behavior of a fly in a room it is natural to start with the case that the fly in in the air, not in the floor, or ceiling, or one of the walls. After investigating the case that the fly is in the air, it is natural to continue the investigation by the case that it is in the floor, or the ceiling, or one of the walls, but not in one of the lines that boundaries two of these planes. After that it is natural to consider the case that the fly is in one of these lines, but not in one of the corners. And the final step is the case that the fly is in one of the corners of the room.

Principle. *Any result in the study of objects of a fixed set O, along with previously obtained (known) results should concern an open set of objects in O (for some problems a stronger requirement: open and dense set of objects).*

In terms of a fly in a room this principle forbids to study the behavior of a fly in the boundary ℓ of the floor F and a wall W unless we know its behavior in $F - \ell$ and $W - \ell$. Here are some mathematical situation violating the given principle:

1. I do not know anything about the geometry of linear operators on $\mathbb{R}^n$ with n distinct eigenvalues and decided to study the case that there are $n-1$ distinct eigenvalues;

2. Let $f(x)$ be a C^∞ function, $f'(x_0) = 0$. I know if x_0 is a local min or local max point if $f''(0) \neq 0$, I know nothing about the case $f''(x_0) = 0$ and decided to attack the case $f''(0) = f'''(x_0) = 0$.

Remark. It is not likely that the given principle is violated by somebody in the way 1. or 2. Nevertheless, in many areas of math one can find a lot of works where this principle is violated. These work include some extremely important for the whole math famous works. This does not mean that the principle is wrong. This means that it is "a way of life", or "religion" which might work for you (your research) well or not, depending on the math you are doing and on the character of results you wish to obtain. No way of life is an obligation, but it is worth to know the potential advantages of main ways of life. I am going to give an idea about the advantages of the given principle (and related principles) for the classification problems of linear algebra and local analysis. In these areas the advantages are so big and clear that following the principle in these problems is almost an obligation for those who wants their research to be respected.

Remark. Applying the principle one should be absolutely sure what is the class O of objects he is working with. The principle assumes that this class is important (interesting, significant), but does not formalize what does it mean. It does not say what is more important: to study the behavior in a room of a fly or a cockroach. A fly should not be confused with cockroach. In the case of cockroach a generic case is the case that it is in the floor, or the ceiling, or in one of the walls.

Remark. In many cases the situation is as follows. One starts with an important class of objects O and comes to conclusion that studying O requires studying another class of different objects $\widetilde{O}$, due to some natural map $F : O \to \widetilde{O}$. In this case the subclass $F(O) \subset \widetilde{O}$ becomes important even if it is a very small subclass of $\widetilde{O}$ and according to the principle any new result for S along with already known ones should concern an open set of objects in $F(O)$ with respect to the topology induced by the topology of O and the map F. Think on Hamiltonial vector fields -

the object occurring in the study of functions on a symplectic space. The principle does not throw away Hamiltonian vector fields!

3. Singularity classes in classification problems

Classification problem is the problem of distinguishing non-equivalent objects of a given class O with respect to a given equivalence relation. The initial data is not only O and an equivalence relation but also a topology in O.

A *singularity class* is a set $S \subset O$ closed with respect to the equivalence: if $a \in S$ then the whole orbit of a (i.e. the set of objects equivalent to a) belongs to S.

Usually O is a manifold (in many cases a vector space), a singularity class S is a smooth or stratified submanifold of O, and one can define the codimension of S. In problems of linear algebra $dimO < \infty$ and $codimS = dimO - dimS$. If S is a stratified submanifold (the union of a finite number of smooth submanifolds satisfying certain "natural" requirements which usually hold) then, by definition, $codimS$ = min of codimensions of the strata. Roughly speaking, it is the number of independent conditions distinguishing S in O. One should not take into account generic conditions (conditions which hold for open sets).

Example 1. Let O be the vector space of linear transformations $T : \mathbb{R}^n \to \mathbb{R}^m$ and $T_1 \sim T_2$ if $U_1 T_1 = T_2 U_2$ for some non-singular linear operators $U_1 : \mathbb{R}^n \to \mathbb{R}^n$ and $U_2 : \mathbb{R}^m \to \mathbb{R}^m$. Let S_r be the set of linear operators of rank r and let $S_{\leq} r$ be the set of linear operators of rank $\leq r$. Then S_r and $S_{\leq r}$ are singularity classes. The class S_r, respectively $S_{\leq r}$, is a smooth, respectively stratified submanifold of O (to see it think about the case $m = n = 2, r = 1$.)

Corank product theorem. In this example,
$codimS_r = codimS_{\leq r} = (m - r)(n - r)$.

In problems of local analysis $dimO = \infty$, but the codimensions of most important singularity classes S is finite. In this case, like in finite dimensional problems, the codimension of S is, roughly speaking, the number of independent conditions distinguishing S in O. A precise definition requires reduction from O to a finite dimensional space of jets (usually singularity classes are distinguished by conditions on k-jet of an object in O, for some k).

Example 2. In the problem of classification of functions $f : \mathbb{R} \to \mathbb{R}$ with respect to the group of diffeomorphisms of the sourse space (f is equivalent to $\tilde{f}$ if $f(\Psi) = \tilde{f}$ for some diffeomorphism Ψ of $\mathbb{R}$) the set of function germs f such that $f'(x_0) = 0$ for some $x_0 \in \mathbb{R}$ is a singularity class of codimension 0; the set of function f such that

$$f'(x_0) = f''(x_0) = \cdots = f^{(k)}(x_0) = 0 \tag{1}$$

for some $x_0 \in \mathbb{R}$ is a singularity class of codimension $k - 1$.

To understand this example note that (1) is a system of k equations for one unknown x_0. The codimension is the difference between the number of equations and the number of unknowns.

Example 3. In the problem of classification of functions $f : \mathbb{R}^2 \to \mathbb{R}$ with respect to the group of diffeomorphisms of the source space $\mathbb{R}^2$, the set of functions

f such that $f'(a) = f''(a) = 0$ for some $a \in \mathbb{R}^2$ is a singularity class of codimension 3. [1]

A singularity class S_1 adjoins a singularity class S_2 (notation: $S_2 \leftarrow S_1$) if S_1 belongs to the closure of S_2 (as submanifolds of O). For example, one has the adjacency the class $S_r \leftarrow S_{r+1}$ in Example 1.

The principle of section 2 can be formulated as follows:

Having a decomposition of O onto singularity classes do not attack a singularity class S before obtaining results for all singularity classes that S adjoins.

Terminology. In many papers "singularity class" = "case". Generic case (see section 1) = a dense singularity class of codimension 0.

4. Normal forms. Modality

A normal form of a single object does not exist. Any normal form always *serves for some set of objects* and speaking about normal forms without mentioning this set is impossible.

A normal form (serving) for a set of objects $S \subset O$ is any set $N \subset S$ such that any $a \in S$ is equivalent to some $b \in N$. A normal form N is called exact if any $a \in S$ is equivalent to one and only one $b \in N$ (or, what is the same, if no two different objects on N are equivalent). Otherwise N is called preliminary normal form (sometimes pre-normal form).

It is natural to construct normal forms serving for singularity classes. If N is an exact normal form serving for some singularity class S, it might be finite, it might be parameterized by a finite number of numerical parameters, and in classification problems of local analysis it might be parameterized by functions. Any of these possibilities holds for some classification problems even if S is a generic singularity class.

The given above principle restricted to the problem of constructing normal forms is as follows:

any new normal form (for some singularity class) along with previously known normal forms (for other singularity classes) should serve for an open singularity class.

Definition. The modality of an object $a \in O$ is the number of numerical parameters of an exact normal form N serving for a sufficiently small neighborhood of a in O. The object a is simple if it has zero modality (i.e. N is a finite set). If a is contained in some singularity class S then the modality of *a within singularity class S* is the number of numerical parameters of an exact normal form N serving for the set $U \cap S$ where U is a sufficiently small neighborhood of a in O.

Distinguishing simple objects in many classification problems and studying the singularity classes {objects of a fixed modality} might be a very important problem, tied with many problems of various areas of math. The classical example is the famous simple A-D-E singularities of functions.

[1]Here $f' = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right)$ and $f'' = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix}$

5. Illustrating examples

Example 5. Let O be the vector space of linear operators $T : \mathbb{R}^n \to \mathbb{R}^m$. In the problem of classifying O with respect to linear transformations of both the source and the target space, the single linear operator

$$x \to A_r x, \quad A_r = \left\{ \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \right\}$$

(I_r is the identity $r \times r$ matrix) is an exact normal form serving for the singularity class of linear operators of rank r. An exact normal form N serving for the whole O consists of $q+1$ linear operators where $q = min(n,m)$, for example

$$N = \{x \to A_i x, \quad i = 0, ..., q\}$$

(A_0 is the zero matrix). The modality of any $T \in O$ is 0.

Example 6. Let O be the vector space of linear operators $T : \mathbb{C}^2 \to \mathbb{C}^2$. In the problem of classifying O with respect to conjugacy (the same linear transformation in the source and the target space), the set

$$N_1 = \left\{ \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{C}, \ \lambda_1 \neq \lambda_2 \right\}$$

is a normal form serving for the singularity class of linear operators with two distinct eigenvalues, the set

$$N_2 = \left\{ \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \quad \lambda \in \mathbb{C} \right\}$$

is a normal form serving for the singularity class of linear operators with one eigenvalue of multiplicity 2. The set $\{N_1 \cup N_2\}$ is a normal form serving for the whole Q. These normal forms are "almost exact" (there is a possibility to replace the eigenvalues); exact normal forms have the same number of parameters. The modality of any linear operator $T \in O$ is equal to 2.

Example 7. Let O be the vector space of real 3×2 matrices. Consider the following equivalence: $A \sim B$ if $TA = B$ for some non-singular 3×3 matrix. The single matrix

$$N_1 = \{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

is an exact normal form serving for the singularity class of rank 2 matrices, and the set

$$N_2 = \left\{ \begin{pmatrix} 1 & a \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad a \in \mathbb{R} \right\}$$

is an exact normal form serving for the singularity class of matrices of rank ≤ 1. The modality of any rank 2 matrix is equal to 0 and the modality of ANY rank ≤ 1 matrix is equal to 1. The modality of any matrix with zero first column *within the singularity class of such matrices* is equal to 0.

6. Principle 2: non-generic case holds in families only

Principle 1 can be explained as follows: all objects are defined up to a small perturbation, therefore results that do not respect a small perturbation are not worth. It seems that following this ideology one should study a generic case only. What for to study singular square matrices if the singularity class $\{A : detA = 0\}$ can be avoided by a small perturbation of A? What for to study function germs $f(x)$ such that $f'(x_0) = f''(x_0)$ for some $x_0 \in \mathbb{R}$ if the singularity class consisting of such functions can be avoided by a small perturbation of $f(x)$?

The answer is as follows: if $a \in O$ is an object belonging to a singularity class S of codimension $d \geq 1$ then we can get rid of S by a small perturbation of a, but we cannot get rid of S by a small perturbation of a generic *d-parameter family of objects which contains a.*

Example 8. Let S be the class of functions $f : \mathbb{R} \to \mathbb{R}$ such that $f(x_0) = f'(x_0) = 0$ for some (non-fixed) $x_0 \in \mathbb{R}$. It has codimension 1 (cf. Example 2). If $f \in S$ then there is an arbitrary small perturbation $\tilde{f}$ of f such that $\tilde{f} \notin S$. For example, $f(x) = x^2 \in S$ and $\tilde{f}(x) = x^2 + \delta$ does not belong to S for any $\delta \neq 0$. On the other hand there are one-parameter families $f_a,\ a \in \mathbb{R}$ of functions such that S is irremovable by their small perturbation. Take, for example, the family $f_a(x) = x^2 + a$ so that $f_0 \in S$. Is it true that if $\widetilde{f_a}(x)$ is a family sufficiently close to $f_a(x)$ then there is a_0 close to 0 such that $\widetilde{f_{a_0}} \in S$, i.e.

$$\widetilde{f_{a_0}}(x_0) = \widetilde{f'_{a_0}}(x_0) = 0 \tag{2}$$

for some x_0 close to 0? The answer is yes provided that "close" holds with respect to sufficiently strong topology in the space of functions. A rough explanation of the answer "yes": (2) is a system of two equations for the same number of unknowns x_0 and a_0.

Example 9. Let S be the class of singular 2×2 matrices. For any $A \in S$ there is an arbitrary small perturbation $\tilde{A} \notin S$. For example,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in S, \quad \tilde{=} \begin{pmatrix} 1 & 0 \\ 0 & \delta \end{pmatrix} \notin S \quad \text{for any } \delta \neq 0.$$

But if we consider a one-parameter family $A_a = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}$ then S cannot be avoided by any smooth perturbation of the family: if $\tilde{A}_a$ is a family sufficiently close to the family A_a, in a sufficiently strong topology, then there exists a_0 close to 0 such that the matrix A_{a_0} is singular.

Example 10. Let S be the class of functions $f : \mathbb{R} \to \mathbb{R}$ such that $f(x_0) = f'(x_0) = f''(x_0) = 0$ for some (non-fixed) $x_0 \in \mathbb{R}$. It has codimension 2 (cf. Example 1). Assume that $f_a(x)$ is a one-parameter family of functions such that $f_{a_*} \in S$ for some a_*. Is it possible to perturb the family $f_a \to \tilde{f}_a$ so that $\tilde{f}_a \notin S$ for any values of the parameter a? The answer is yes. Take for example the family $f_a(x) = x^3 + ax$ so that $f_0 \in S$. Consider the family $\tilde{f}_a(x) = x^3 + ax + \epsilon$. If $\epsilon \neq 0$ then $\tilde{f}_a \notin S$ for any $a \in \mathbb{R}$.

On the other hand, the class S is irremovable in 2-parameter families. Consider the 2-parameter family $f_{a,b}(x) = x^3 + a + bx$. If one uses a strong enough topology, the following can be proved: if $\tilde{f}_{a,b}(x)$ is a family sufficiently close to $f_{a,b}(x)$ then there exist (a_0, b_0) close to $(0,0)$ such that $f_{a_0,b_0} \in S$. A rough explanation: the system

$$f_{a,b}(x_0) = f'_{a,b}(x_0) = f''_{a,b}(x_0) = 0$$

consists of three equations for the same number of unknowns x_0, a, b.

Example 11. The class S of 2×4 matrices of rank 1 has codimension 3. It can be avoided by a small perturbation of a family of matrices with ≤ 2 parameters and it is irremovable (in the same sense as in examples above) in families with 3 parameters.

Principle 2. Along with studying individual objects of a singularity class S of codimension d one should study families of objects such that S is irremovable by a small perturbation of a family. Such families have $\geq d$ parameters.

7. Versal deformations

A deformation of an object $a \in O$ is any family a_ϵ such that $a_0 = a$ and a_ϵ depends smoothly on ϵ. Here $\epsilon \in \mathbb{R}^m$ and the number of parameters m can be arbitrary. We assume that ϵ is small meaning that ϵ varies in as small as we wish neighborhood of $0 \in \mathbb{R}^m$. In view of Principle 2 it is worth to ask about a normal form serving for *any* deformation of a fixed $a \in O$.

It seems that if we a normal form serving for O or for an open singularity class containing a then we have an answer.

Example 12. Consider the vector space O of linear operators $T : \mathbb{R}^2 \to \mathbb{R}^3$ and consider the problem of classifying O with respect to linear transformations of both the source and the target space. Consider the operator A with the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ (here and below we fix a basis of $\mathbb{R}^2$ and a basis of $\mathbb{R}^3$). Any linear operator in any deformation of A is equivalent to an operator with one of the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{30}$$

and one can say that these two matrices is a normal form serving for all deformations A_ϵ of A. This is so, but this normal form has a big disadvantage: most deformations A_ϵ cannot be reduced to this normal form by linear transformations of the source and the target vector spaces depending smoothly (or even continuously) on ϵ. How to construct a

(*) normal form such that any linear operator of any deformation A_ϵ can be reduced to some linear operator of this normal form by transformations of the source and target vector spaces depending smoothly on ϵ?

It is not hard to prove that one of examples of such normal form is the set of deformations

$$\begin{pmatrix} 1 & 0 \\ 0 & f_1(\epsilon) \\ 0 & f_2(\epsilon) \end{pmatrix} \tag{4}$$

with arbitrary functions $f_1(\epsilon), f_2(\epsilon)$ vanishing at $\epsilon = 0$, and this normal form cannot be simplified in the sense that any normal form satisfying (*) will contain two arbitrary functions $f_1(\epsilon), f_2(\epsilon)$ vanishing at $\epsilon = 0$.

Note that die to (*) normal form (4) is much more informative than (3) . It shows, for example, how the image of A_ϵ depends on ϵ.

Replace in (4) the function $f_1(\epsilon)$ by δ_1 and $f_2(\epsilon)$ by δ_2 (certainly other letters for small parameters can be used). We obtain

$$\begin{pmatrix} 1 & 0 \\ 0 & \delta_1 \\ 0 & \delta_2 \end{pmatrix}.$$

This 2-parameter family is called a mini-versal deformation of the matrix A (in our classification problem).

Definition. Assume that the equivalence of objects in O is induced by the action of a group G. A versal deformation of $a \in O$ is a smooth deformation $A_{\delta_1,\dots,\delta_s}$ of a such that for any smooth deformation a_ϵ of a, with any number of parameters ($\epsilon \in \mathbb{R}^m$, any m), there exists a smooth deformation $g_\epsilon \in G$ of the identity element of g ($g_0 = id$) such that $g_\epsilon.a_\epsilon = A_{f_1(\epsilon),\dots,f_s(\epsilon)}$ for some smooth functions $f_1(\epsilon), ..., f_s(\epsilon)$. A versal deformation is called mini-versal if the number s of its parameters is minimal possible.

Of course, the word "smooth" should be formalized, especially when $dimO = dimG = \infty$ - the case of classification problems of local analysis. Certainly, O and G are subjects to certain restrictions (which hold in most valuable classification problems).

Example 13(exercise). In the problem of classification of 2×2 matrices with respect to similarity, one of the mini-versal deformations of the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ is the deformation $\begin{pmatrix} 1+\delta_1 & 0 \\ 0 & 2+\delta_2 \end{pmatrix}$.

It is more difficult to construct a mini-versal deformation of the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. In the lecture I said that one of the mini-versal deformations is

$$\begin{pmatrix} 1+\delta_1 & 1 \\ 0 & 1+\delta_2 \end{pmatrix}. \tag{5}$$

I made a mistake: it is wrong.

Example 14 (exercise). Consider the one-parameter deformation $A_\epsilon = \begin{pmatrix} 1 & 1 \\ \epsilon & 1 \end{pmatrix}$ and prove that there are no smooth functions $T_{ij}(\epsilon), f_1(\epsilon), f_2(\epsilon)$ such that

$$T(0) = I,\ f_1(0) = f_2(0) = 0,\ \ T^{-1}(\epsilon)A_\epsilon T(\epsilon) = \begin{pmatrix} 1+f_1(\epsilon) & 1 \\ 0 & 1+f_2(\epsilon) \end{pmatrix},$$

$$T(\epsilon) = \begin{pmatrix} T_{11}(\epsilon) & T_{12}(\epsilon) \\ T_{21}(\epsilon) & T_{22}(\epsilon) \end{pmatrix}$$

Consequently (5) is NOT a versal deformation of the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Example 15 (exercise). Prove that in the problem of classifying 2×2 matrices with respect to similarity, each of the families

$$\begin{pmatrix} 1 & 1 \\ \delta_1 & 1+\delta_2 \end{pmatrix}, \quad \begin{pmatrix} 1+\delta_1 & 1 \\ \delta_2 & 1+\delta_1 \end{pmatrix}$$

is a mini-versal deformation of the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Theorem. In classification problems of linear algebra and in certain classification problems of local analysis (including local classification of maps from $\mathbb{R}^n \to \mathbb{R}^m$) the number of parameters of any mini-versal deformation of an object a in a singularity class S is the codimension of S plus the modality of a within S.

Exercise. Check this theorem for all examples above and try to understand it (rather than to prove).

Example 16 (exercise). Let $S \subset Mat(2,2)$ be the singularity class of diagonalizable matrices with one eigenvalue of multiplicity 2 (equivalence: similarity of matrices).

a) find $codimS$

b) find the modality of any $A \in S$ within S

c) use the given theorem to find the number of parameters of any mini-versal deformation of A

d) find a mini-versal deformation of the matrix I

Answers: a) 3 b) 1 c) $3+1=4$ d) $\begin{pmatrix} 1+\delta_1 & \delta_2 \\ \delta_3 & 1+\delta_4 \end{pmatrix}$.

8. Illustrating example: the Gauss method for solving systems of linear equations

Given a system of linear equations $Ax = b$ where $A \in Mat(m,n), b \in \mathbb{R}^m, x \in \mathbb{R}^n$ associate to it the $m \times (n+1)$ matrix (A, b) whose first n columns are the columns of A and the last column is the vector b. The Gauss method is the solution of the problem of classifying matrices (A, b) with respect to the following equivalence: $C = (A, b)$ is equivalent to $\tilde{C} = (\tilde{A}, \tilde{b})$ if $TC = \tilde{C}$ for some non-singular $m \times m$ matrix T. Two equivalent matrices $C, \tilde{C}$ correspond to linear systems with the same general solution. The Gauss method is the reduction of C to a certain normal form. If C belongs to this normal form then the corresponding linear system can be solved immediately.

Consider the simplest case $m = n = 2$, i.e. solving linear systems of two equations for two unknowns. In this case the Gauss method is the reduction to a certain normal form of 2×3 matrices C with respect to the given above equivalence. For the generic singularity class consisting of matrices with non-zero first column this normal form is as follows:

$$\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix}, \quad \text{(unique solution)}$$
$$\begin{pmatrix} 1 & a & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{(no solutions)}$$
$$\begin{pmatrix} 1 & a & b \\ 0 & 0 & 0 \end{pmatrix} \quad \text{(infinitely many solutions).}$$

It is easy to obtain mini-versal deformations:

$$\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & a+\delta_1 \\ 0 & 1 & b+\delta_2 \end{pmatrix}. \quad \text{Solution: } x_1 = a+\delta_1,\ x_2 = b+\delta_2;$$
$$\begin{pmatrix} 1 & a & 0 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & a+\delta_1 & 0 \\ 0 & \delta_2 & 1 \end{pmatrix}. \quad \text{Solution: } x_1 = -\tfrac{a+\delta_1}{\delta_2},\ x_2 = \tfrac{1}{\delta_2} \quad (\text{for } \delta_2 \neq 0)$$
$$\begin{pmatrix} 1 & a & b \\ 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & a+\delta_1 & b+\delta_2 \\ 0 & \delta_3 & \delta_4 \end{pmatrix}. \quad \text{Solution } x_1 = b+\delta_2 - \tfrac{(a+\delta_1)\delta_4}{\delta_3},\ x_2 = \tfrac{\delta_4}{\delta_3}$$
$$(\text{for } \delta_3 \neq 0)$$

If the parameters vary in a certain open dense set U (i.e. in a generic case in the space of parameters) the solution is unique and the mini-versal deformations show the qualitative dependence of the solution on the parameters varying in U.

Exercise. Construct a normal form serving for the whole space of 2×3 matrices (including matrices with zero first column) and corresponding mini-versal deformations. Use the mini-versal deformations to determine the qualitative difference between the following two systems without solutions:

$$\begin{pmatrix} 1 & a \\ 0 & 0 \end{pmatrix} x = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$