1 Introduction

These are the notes of the 42-hours course "Ordinary Differential Equations" ("Equazioni Differenziali Ordinarie") that I am going to deliver for the "Laurea in Matematica" of the University of Trento. I have decided to write them in English (even if the course will be delivered in Italian), since in the future they may be also suitable for the "Laurea Magistralis in Mathematics", whose official language is English.

To profitably read these notes, some background on Analysis, Geometry and Linear Algebra is needed (roughly speaking: the subject of the Analysis and Geometry courses of the first two years of the degree in Mathematics). Other advanced concepts will be somewhere used: convergence of functions, topology for functions spaces, complete metric spaces, advanced linear algebra. A familiarity with such concepts is certainly welcome. However, in the Appendix some of the requested notions and results are briefly treated.

The mathematical notations in these notes are the standard ones. In particular $\mathbb{R}$, $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{C}$ respectively stand for the sets of real, natural, rational and complex numbers. Moreover, if $n \in \mathbb{N} \setminus \{0\}$, then $\mathbb{R}^n$ is, as usual, the n-dimensional space $\mathbb{R} \times \ldots \times \mathbb{R}$ i.e. the cartesian product of $\mathbb{R}$ n-times. The points (or vectors) of $\mathbb{R}^n$, when thought as string of coordinates, will be usually written as a line, even if, when a $m \times n$ matrix is applying to them, they must be think as column. With the notation $[a, b]$ we will mean the interval of real numbers $\{x \in \mathbb{R} | a \leq x \leq b\}$, that is the closed interval containing its extreme points. In the same way $]a, b[$ will denote the open interval without extreme points $\{x \in \mathbb{R} | a < x < b\}$[1], $]a, b]$ the semi-open interval $\{x \in \mathbb{R} | a < x \leq b\}$ and $[a, b[$ the semi-open interval $\{x \in \mathbb{R} | a \leq x < b\}$[2].

The time-derivative will be usually denoted by "y' ", and very rarely by "$\dot{y}$".

The Euclidean norm in $\mathbb{R}^n$, with $n > 1$, will be denoted by $\|\cdot\|_{\mathbb{R}^n}$ or, if there is no ambiguity, simply by $\|\cdot\|$. The absolute value of $\mathbb{R}$ is denoted by $|\cdot|$. When $x \in \mathbb{R}^n$, with $n > 1$, and when $r > 0$, with the notation $B_{\mathbb{R}^n}(x, r)$ we will denote the open ball centered in x with radius r

$$B_{\mathbb{R}^n}(x, r) = \{z \in \mathbb{R}^n | \|z - x\|_{\mathbb{R}^n} < r\}.$$

Also in this case, if no ambiguity arises, we may use the simpler notation $B(x, r)$.

If A is a subset of $\mathbb{R}^n$, by $\overline{A}$ we will denote its closure.

In these notes the formulas will be enumerated by $(x.y)$ where x is the number of the section (independently from the number of the subsection) and y is the running number of the formula inside the section. Moreover, the statements will be labeled by "S $x.y$" where "S" is the type of the statement (Theorem, Proposition, Lemma, Corollary, Definition, Remark, Example), x is the number of the section (independently from the number of the subsection), and y is the running number of the statement inside the section (independently from the type of the statement).

The symbol "□" will mean the end of a proof.

[1] Here, we of course permit $a = -\infty$ as well as $b = +\infty$.

[2] In the last two cases we, respectively, permit $a = -\infty$ and $b = +\infty$.

Some possible references are[3]

- G. De Marco: *Analisi Due -seconda parte*, Decibel-Zanichelli, Padova 1993.
- C.D. Pagani - S. Salsa: *Analisi Matematica, Volume 2*, Masson, Milano 1991.
- L.C. Piccinini, G. Stampacchia, G. Vidossich: *Ordinary Differential Equations in* $\mathbb{R}^n$, Springer-Verlag, New York 1984.
- W. Walter: *Ordinary Differential Equations*, Springer-Verlag, New York 1998.

Please feel free to point out to me the mathematical as well as the english mistakes, which are for sure present in the following pages.

Let's start!

1.1 Motivating examples

The theory of ordinary differential equations is one of the most powerful method that humans have invented/discovered[4] and continuously improved for describing the natural phenomena whose investigation is fundamental for the progress of humanity. But its power is not limited to the "natural phenomena" (physical, biological, chemical etc.), it is also fundamental for the study and the construction of mechanical systems (engineering) as well as for the study and the prediction of the economical/social behavior of our real world.

An ordinary differential equations is a functional equation which involves an unknown function and its derivatives. The term "ordinary" means that the unknown is a function of a single real variable and hence all the derivatives are "ordinary derivatives".

A natural extension of the theory of ordinary differential equations is the theory of partial differential equations[5], which is certainly more suitable for describing those phenomena whose space-dependence is not negligible. However, most of the results about partial differential equations were not be obtainable without a good theory for the ordinary differential equations.

Since the unknown function depends only on a real variable, it is natural to give it the meaning of <u>time</u>, denoting it by $t \in \mathbb{R}$, and to interpret the solution as the <u>evolution</u> of the system under study. Here are some examples in this sense.

[3] The first two have strongly inspired the present notes, the second two may be suggested for deeper readings.

[4] Is Mathematics invented or is it discovered? We do not enter in such a diatribe. We leave it to philosophers.

[5] *A partial differential equation* is a functional equation which involves an unknown function and its derivatives. The term "partial" means that the unknown depends on many real variables and hence the derivatives in the equation are "partial derivatives".

Example 1.1 Capital management. The evolution of the amount of a capital at disposal is represented by a time-dependent function $K(\cdot)$. At every instant[6], there is a fraction cK which is re-invested with an instantaneous interest rate given by i, and there is also a fraction dK which is spent without earning. Here c, d, i are all real numbers between 0 and 1 and $c + d \leq 1$. The evolution law for the capital is then given by the equation

$$K'(t) = (ic - d)K(t), \tag{1.1}$$

which means that K *instantaneously tends to increase for the capitalization of the re-invested quantity* cK*, and instantaneously tends to decrease for the spent quantity* dK. Even if this is a very simple model, it is obvious that the possibility of computing the capital evolution $K(\cdot)$, i.e of solving the equation (1.1), is extremely important for the management of the capital. For instance, one may be interested in suitably choosing the coefficients c and d[7] in order to get a desired performance of the capital without too much reducing the expendable amount. To this end, the capability of solving the equation is mandatory. Here, a solution is a one-time derivable real-valued function K, defined on a suitable interval $I \subseteq \mathbb{R}$, $K : I \to \mathbb{R}$.

Example 1.2 Falling with the parachute. A body is falling hanged to its parachute. Denoting by g the gravity acceleration and by $\beta > 0$ the viscosity coefficient produced by the parachute, the law of the motion is, in a upward oriented one-dimensional framework,

$$x''(t) = -g - \beta x'(t). \tag{1.2}$$

This means that the time-law of the fall (i.e. the time-dependent function $x(\cdot)$) must solve the functional equation (1.2), which says: *the acceleration of the falling body is given by the downward gravity acceleration plus a term which depends on the velocity of the fall and on a viscosity coefficient.* This last term is responsible of the fall's safety: bigger is β slower is the fall[8]. With air's viscosity fixed, the coefficient β depends only on the shape of the parachute. Hence, one may be interested in calculating a suitable[9] coefficient β and then construct a corresponding parachute. It is obvious that the "suitableness" of β may be tested only if we know the corresponding evolution x, that is only if we can solve the equation (1.2) for all fixed value of β. A solution is then a two-times derivable real-valued function x, defined on a suitable interval $I \subseteq \mathbb{R}$, $x : I \to \mathbb{R}$.

Example 1.3 Filling a reservoir. A natural water reservoir is modeled by a bidimensional rectangular box $[a, b] \times [0, H]$. Let us suppose that the reservoir is filled in by a water source whose rate of introduction of water is constantly equal to $c > 0$ (volume of

[6] Let us suppose that the model is based on a "continuous time" instead of a "discrete time" (for instance:day by day, week by week...) as may be more natural to assume. The "continuity" of the time may be reasonable if we are looking to the evolution of the capital in a long period: ten, twenty years.

[7] Unfortunately, the most important coefficient in (1.1), that is i, is not at disposal of the investor, but it is decided by the bank. And also changing bank is not helpful.

[8] Note that, when the body is falling, x' is negative and hence $-\beta x'$ is positive

[9] "Suitable" could mean: safely reach the ground without spending too much time.

water per unit time). Let us suppose that the vertical layers of the reservoir have some degree of porosity. This means that, at every instant t, an amount of water exits through the points (a, h), (b, h) with a rate that is proportional to the porosity and to the quantity of water over the point (the pressure). In particular, let us denote by $u(t)$ the level of water inside the reservoir at the time t and let us suppose that the porosity depends on the height[10]. This means that there exists a function $g : [0, h] \to [0, 1]$ such that the rate of exit through (a, h) and (b, h) at time t is equal to zero if $u(t) < h$, otherwise it is given by $g(h)(u(t) - h)$.[11] Hence, the rate at the time t of the total volume of water that exits through the vertical layers is given by

$$\int_0^{u(t)} g(h)(u(t) - h)dh.$$

Hence, the instantaneous variation of the level u, that is its time derivative, is given by

$$u'(t) = \frac{c}{b-a} - \frac{1}{b-a}\int_0^{u(t)} g(h)(u(t) - h)dh.$$

For instance, if

$$g(h) = \frac{h}{H},$$

which means that the more permeable soils are on the surface, we then get the equation

$$u'(t) = \frac{c}{b-a} - \frac{u^3(t)}{6(b-a)H}. \tag{1.3}$$

If the source supplies water with a rate which is not constant but it depends on time, let us say $c(t)$, then the equation is

$$u'(t) = \frac{c(t)}{b-a} - \frac{u^3(t)}{6(b-a)H}. \tag{1.4}$$

Being able to calculate the solution u of (1.3) (or of (1.4)) permits to predict whether (and possibly at which time) the reservoir will become empty, or filled up, or whether its level will converge to a stable value. Here a solution is a one-time derivable real-valued function.

Example 1.4 The prey-predator model of Lotka-Volterra. Two species of animals, X and Y, occupy a certain region and interact. Let us denote by $x(t) \geq 0$ and by $y(t) \geq 0$ the number of present individuals at time t for both species respectively. Let us suppose that

[10] Such a situation is common in the application where, for deep reservoirs, different types of soils are stratified along the wall of the reservoir.

[11] Note that the point is not permeable if $g(h) = 0$, and hence nothing exit through that point; it is completely permeable if $g(h) = 1$, and hence if $u(t) > h$, then through that point the water exits with rate given by $(u(t) - h)$. If $0 < g(h) < 1$ all the intermediate cases may hold.

the species Y is given by predators whose preys are exactly the individuals X. Moreover, we suppose that the relative rate of increasing for the preys (i.e x'/x) is constant and positive when there are no predators ($y = 0$), and instead linearly decreases as function of y in case of presence of predators ($y > 0$). On the other hand, we suppose that the relative rate of increasing for the predators (i.e. y'/y) is constant and negative when there are no preys ($x = 0$)[12], and instead is linearly increasing as function of x in case of presence of preys ($x > 0$). Hence we have the following system

$$\begin{cases} \dfrac{x'}{x} = \alpha - \beta y \\ \dfrac{y'}{y} = -\gamma + \delta x \end{cases},$$

where $\alpha, \beta, \gamma, \delta$ are positive constants. The system can also be written as

$$\begin{cases} x' = x(\alpha - \beta y) \\ y' = y(-\gamma + \delta x) \end{cases} \tag{1.5}$$

Solving (1.5) may permit to study the evolution of the two species, which is certainly important from many points of view[13]. A solution of (1.5) is a one-time derivable vectorial function $t \mapsto (x(t), y(t)) \in \mathbb{R}^2$.

Up to now, we have considered model problems where the variable of the unknown function has the meaning of time, and the solution the meaning of evolution. However, this not the only case (even if it is a natural framework). Next two examples show cases where the real variable of the unknown function does not have the meaning of time, but rather the meaning of <u>space</u>[14] [15].

Example 1.5 <u>The catenary curve.</u> A homogeneous chain is hanged to a vertical wall by its extremes on two points, not on the same vertical line, and it is subject only to the gravity force. Which is the shape attained by the chain? Let us suppose that the shape of the chain is given by the graph of a function $y : [a, b] \to \mathbb{R},\ x \mapsto y(x)$, where a and b are the abscissas of the two hanging points. On every piece of arch of the chain, the resultant of all the applied forces must be zero, since the chain is in equilibrium. Such forces are: 1) the total weight of the piece of arch, 2) the force exerted on the right extremum by the remaining right part of the chain, 3) the force exerted on the left extremum by the remaining left part of the chain. The first force is vertical and downward, the other two are tangential. Let our piece of arch be the

[12]They have nothing to eat.

[13]Nowadays we can in particular say "from an ecological point of view".

[14]However, in the sequel, we will often adopt the point of view of "time" and "evolution". It is obvious that, from an analytical point of view, the meaning given to the variable (time, space, what else...) and the name given to it (t, x, p...) is completely meaningless.

[15]The second of the next examples (Optimal control) starts from a problem of "evolution" but in the ordinary differential equation (1.7) the variable p has the meaning of space: the starting point of the evolution.

part of the graph over the subinterval $[x_l, x_r] \subseteq [a, b]$. We write our three forces by their horizontal and vertical components in the following way: 1) $\mathbf{P} = (0, -p)$ distributed on all the piece of arch, 2) $\mathbf{T}^r(x_r, y(x_r)) = (T_1^r(x_r, y(x_r)), T_2^r(x_r, y(x_r)))$, 3) $\mathbf{T}^l(x_l, y(x_l)) = (T_1^l(x_l, y(x_l)), T_2^l(x_l, y(x_l)))$. Since we must have

$$\mathbf{P} + \mathbf{T}^r(x_r, y(x_r)) + \mathbf{T}^l(x_l, y(x_l)) = (0, 0),$$

we then deduce that the modules of the horizontal component of $\mathbf{T}^r(x_r, y(x_r))$ and of $\mathbf{T}^l(x_l, y(x_l))$ are equal. By the arbitrariness of x_l and x_r, we deduce that such a modulus is constant, let us denote it by $c > 0$. Now, let $x_v \in [a, b]$ be a point of minimum for y (i.e. a point of minimum height for the chain). On $(x_v, y(x_v))$ the tangent is then horizontal, and, if we repeat the previous argument on the interval $[x_v, x_r]$, we have that $\mathbf{T}^l(x_v, y(x_v))$ has null second component (since it is tangent). Hence, the vertical weight must be balanced only by the vertical component of $\mathbf{T}^r(x_r, y(x_r))$. Let us denote by g the modulus of the gravity acceleration, and by μ the constant linear mass-density of the chain. Hence the weight of the arch over the interval $[x_v, x_r]$ is given by[16]

$$p = \int_{x_v}^{x_r} g\mu\sqrt{1 + (y'(x))^2}dx.$$

We then get

$$T_2^r(x_r, y_(r)) = \int_{x_v}^{x_r} g\mu\sqrt{1 + (y'(x))^2}dx.$$

Since the ratio $T_2^r(x_r, y_(r))/T_1^r(x_r, y_(r)) = T_2^r(x_r, y_(r))/c$ is the angular coefficient of the graph of y in $(x_r, y(x_r))$, that is $y'(x_r)$, we also get, for the arbitrariness of $x_r \geq x_v$ and repeating similar consideration for points to the left of x_v,

$$y'(x) = \frac{g\mu}{c}\int_{x_v}^{x} \sqrt{1 + (y'(\xi))^2}d\xi, \quad \forall x \in]a, b[.$$

Differentiating, we finally obtain

$$y''(x) = \frac{g\mu}{c}\sqrt{1 + (y'(x))^2} \quad \forall x \in]a, b[. \tag{1.6}$$

Being able to calculate the solution of (1.6) permits to know the shape of the chain[17].

[16]It is the curvilinear integral of the infinitesimal weight $g\mu ds$, where $ds = \sqrt{1 + (y'(x))^2}dx$ is the infinitesimal element of length of the arch.

[17]Actually, to exactly know the shape of the chain, we need some other information: the heights of the hanging points and the total length of the chain, otherwise many solutions are possible. But this is a common feature. Also in previous examples we usually need some other information as, for instance, the value of the solution in a fixed instant.

Example 1.6 Optimal control. A material point is constrained to move, without friction, on a one-dimensional guide. On the guide there is a system of coordinates, let us say $p \in \mathbb{R}$. The material point has to reach a fixed point positioned on $\overline{p}$, and it has several choices for moving: one per every value of the parameters $a \in [-1, 1]$. That is, at every instant $t \geq 0$, for every choice of a, it moves with instantaneous velocity equal to a. However, every such a choice has a cost, which depends on the actual position p of the point on the guide and on the parameter a, via an instantaneous cost function $\ell(p, a)$. The goal is to reach the target point $\overline{p}$ using a suitable moving strategy $a(t) \in [-1, 1]$ in order to minimize the following quantity

$$t^* + \int_0^{t^*} \ell(p(t), a(t))dt,$$

where t^* is the reaching time of the target point and $p(\cdot)$ is the evolution of our point. In other words, using a suitable moving strategy $a(\cdot)$, we want to reach the target point trying to minimize a "combination" of the spent time and the total cost given by ℓ[18]. For every starting point p, we can consider the "optimal function" $V(p)$ which is the optimum (i.e. the minimum cost) that we can get starting from p. Under suitable hypotheses[19], V solves the following equation

$$\sup_{a \in [-1,1]} \{-V'(p)a - \ell(p, a)\} = 1. \tag{1.7}$$

Being able to calculate V from (1.7) may permit to get useful information on the minimization problem, for instance on how to construct an optimal strategy $a(\cdot)$[20].

1.2 Notations, definitions and further considerations

An ordinary differential equation is an expression of the following type:

$$F(t, y^{(n)}, y^{(n-1)}, \dots, y', y) = 0, \tag{1.8}$$

where $F : \mathcal{D} \subseteq \mathbb{R} \times \mathbb{R}^k \times \cdots \times \mathbb{R}^k = \mathbb{R} \times \left(\mathbb{R}^k\right)^{n+1} \to \mathbb{R}$ is a function; $t \in \mathbb{R}$ is a scalar parameter; y is the unknown function, which is a function of t and takes values in $\mathbb{R}^k$ with $k \in \mathbb{N} \setminus \{0\}$; y' is the first derivative of y and, for every i, $y^{(i)}$ is the i-th derivative y. The highest order of derivation which occurs in the equation is said the order of the equation. If $k = 1$, then the equation is said a scalar equation.

Solving (1.8) means to find an interval $I \subseteq \mathbb{R}$ and a n-times derivable function $y : I \to \mathbb{R}^k$ such that for every $t \in I$

[18] The idea is that using values of a which give high velocities, and then reduce the spent time a lot, is probably not convenient from the point of view of ℓ: such velocities may be expensive. Hence a suitable combination of high and cheap velocities is needed.

[19] Actually, such "suitable hypotheses" are very delicate.

[20] However, in the real application, the evolution of the point p is not one-dimensional, but it is an evolution in $\mathbb{R}^n$, and hence the equation (1.7) satisfied by the optimum V is a partial differential equation, with V' replaced by the gradient ∇V.

$$(t, y^{(n)}(t), y^{(n-1)}(t), \dots, y'(t), y(t)) \in \mathcal{D},$$

and also that

$$F(t, y^{(n)}(t), y^{(n-1)}(t), \dots, y'(t), y(t)) = 0.$$

The function y is said a <u>solution of the ordinary differential equation</u>.

Let us note that a solution is a function from an interval of the real line to $\mathbb{R}^k$. Then it is a parametrization of a curve in $\mathbb{R}^k$, and the law $t \mapsto y(t)$ is the time-law of running such a curve. For this reason, a solution of an ordinary differential equation is sometimes called a <u>trajectory</u>.

An ordinary differential equation may have infinitely many solutions, or finitely many solutions or even no solutions at all. We define the <u>general integral of the equation</u> as the following set (which may be infinite, finite, a singleton or empty)

$$\mathcal{I} := \left\{ y : I \to \mathbb{R}^k \middle| I \subseteq \mathbb{R} \text{ is an interval and } y \text{ is solution of the equation} \right\} \tag{1.9}$$

In other words the general integral is the set of all solutions of the equation, everyone defined on its interval of definition.

The equation (1.8) is in the so-called <u>non-normal form</u>, that is the highest order derivative (the one of order n, in our case) is not "privileged", that is it is not "isolated", it is not "outside" from the function F. On the contrary we say that an ordinary differential equation is <u>in normal form</u> if it is of the following type

$$y^{(n)} = f(t, y^{(n-1)}, \dots, y', y), \tag{1.10}$$

where $f : \mathcal{D} \subseteq \mathbb{R} \times \mathbb{R}^k \cdots \times \mathbb{R}^k \to \mathbb{R}^k$. In particular, let us note that the co-domain of f is now $\mathbb{R}^k$ (and not anymore $\mathbb{R}$), and that the domain is contained in $\mathbb{R} \times \left(\mathbb{R}^k\right)^n$ (and not anymore $\mathbb{R} \times \left(\mathbb{R}^k\right)^{n+1}$). Hence, we have a <u>system of k scalar differential equations of n order</u>[21]. Indeed, denoting $y = (y_1, \dots, y_k)$ and $f = (f_1, \dots, f_k)$ by their components, we get

$$\begin{cases} y_1^{(n)} = f_1 \left(t, (y_1^{(n-1)}, \dots, y_k^{(n-1)}), \dots, (y_1', \dots, y_k'), (y_1, \dots, y_k) \right) \\ y_2^{(n)} = f_2 \left(t, (y_1^{(n-1)}, \dots, y_k^{(n-1)}), \dots, (y_1', \dots, y_k'), (y_1, \dots, y_k) \right) \\ \cdots \qquad \qquad \cdots \\ y_k^{(n)} = f_k \left(t, (y_1^{(n-1)}, \dots, y_k^{(n-1)}), \dots, (y_1', \dots, y_k'), (y_1, \dots, y_k) \right) \end{cases}$$

It is evident that all equations (systems) in normal form may be written in a non-normal form, for instance by

[21] Actually, systems of s non-normal equations as (1.8) in m unknowns may also be considered. However, we will not discuss such a situation.

$$F(t, y^{(n)}, \dots, y', y) := \left\| y^{(n)} - f(t, y^{(n-1)}, \dots, y', y) \right\|_{\mathbb{R}^k}.$$

On the contrary not all equations in non-normal form may be written in normal form. This depends on the solvability of the algebraic equation $F = 0$ with respect to its second entry $y^{(n)}$. For instance, the second order scalar equation

$$F(t, y'', y', y) = (y'')^2 - 1 = 0,$$

where F has domain $\mathcal{D} = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ cannot be "globally" written in normal form, that is there is not a function $f(t, y', y)$ such that

$$F(t, y'', y', y) = 0 \Leftrightarrow y'' = f(t, y', y),$$

for every values $(t, y''(t), y'(t), y(t)) \in \mathcal{D}$. For example, the functions

$$y_1 : t \to \frac{t^2}{2}, \qquad y_2 : t \to -\frac{t^2}{2}$$

satisfy the equation $F(t, y'', y', y) = 0$, but they cannot satisfy $y'' = f(t, y', y)$ with the same f because otherwise, for $t = 0$, we should have

$$1 = f(0, 0, 0) = -1.$$

Hence, the property of being in normal form or in non-normal form does not depend on how we write the equation, but it is an intrinsic feature of the equation itself.

In general, the normal form equations are simpler to study.

An ordinary differential equation of the general form (1.8)[22] is linear homogeneous if it is linear in the unknown function y and its derivatives. That is if F is linear with respect to its second $n + 1$ components. In other words if for every n-times differentiable functions $u, v : I \to \mathbb{R}^k$, for every $t \in I$, and for every scalars α, β, we have

$$\begin{aligned} &F(t, (\alpha u + \beta v)^{(n)}(t), \dots, (\alpha u + \beta v)'(t), (\alpha u + \beta v)(t)) \\ &= \alpha F(t, u^{(n)}(t), \dots, u'(t), u(t)) + \beta F(t, v^{(n)}(t), \dots v'(t), v(t)). \end{aligned}$$

An ordinary differential equation is said to be a linear nonhomogeneous equation if it is of the form

$$F(t, y^{(n)}, \dots, y', y) = g(t)$$

with F linear as before.

An ordinary differential equation is said to be a autonomous equation if it does not explicitly depend on the scalar variable $t \in \mathbb{R}$. Again referring to (1.8), we must have, for the non-normal form

[22] A similar definition obviously holds for equation in normal form, when we look to the linearity of the function f with respect to its second n components.

$$F(y^{(n)}, \dots, y', y) = 0,$$

with $F : \mathcal{D} \subseteq \left(\mathbb{R}^k\right)^{n+1} \to \mathbb{R}$, and, for the normal form

$$y^{(n)} = f(y^{(n-1)}, \dots, y', y),$$

with $f : \mathcal{D} \subseteq \left(\mathbb{R}^k\right)^n \to \mathbb{R}^k$. If the equation explicitly depends on t, we then speak of nonautonomous equation.

The systems of first-order equations in normal form, $y' = f(t, y)$, have a particular importance. Indeed, they are suitable for describing many evolutionary applied models. Moreover, every n-order scalar equation in normal form may be written as a first-order system of n scalar equations. Indeed, if we have the equation

$$y^{(n)} = g(t, y^{(n-1)}, \dots, y', y), \tag{1.11}$$

with $g : \mathcal{D} \subseteq \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, then we can define $Y_0 = y, Y_1 = y', \dots, Y_{n-1} = y^{(n-1)}$, $Y = (Y_0, \dots, Y_{n-1}) \in \mathbb{R}^n$, and write the system

$$\begin{cases} Y_0' = Y_1 \\ Y_1' = Y_2 \\ \dots \\ Y_{n-2}' = Y_{n-1} \\ Y_{n-1}' = g(t, Y_{n-1}, \dots, Y_1, Y_0) \end{cases} \tag{1.12}$$

It is evident that $y : I \to \mathbb{R}$ is a solution of (1.11) if and only if y is the first component of $Y : I \to \mathbb{R}^n$ with Y solution of (1.12). If we define

$$f(t, Y_0, \dots, Y_{n-1}) = (Y_1, Y_2, \dots, Y_{n-1}, g(t, Y_{n-1}, \dots, Y_1, Y_0)),$$

we then may write the system (1.12) as $Y' = f(t, Y)$.

Concerning the "evolutionary" feature of first-order systems, we will often use the interpretation of the solutions $y : I \to \mathbb{R}^k$ as trajectories (or curves) in $\mathbb{R}^k$, where y is the parametrization and I is the set of parameters (thought as "time"). For the particular case of first-order systems, the equality $y'(t) = f(t, y(t))$ means that, for every $t \in I$ and for every point $x = y(t)$ of the trajectory, the tangent vector to the trajectory itself is exactly given by $f(t, x) = f(t, y(t))$. In other words, if a particle is moving around $\mathbb{R}^k$ with the condition that, at any instant t, its vectorial velocity is given by $f(t, x)$, where x is the position of the particle at the time t, then the particle is necessarily moving along a trajectory given by a solution of the system. That is, if, for every time t and every position x, we assign the vectorial velocity of a motion by the law $v = f(t, x)$, then the motion must be along a trajectory solution of the system. In this setting, the function f is sometimes called dynamics and $\mathbb{R}^k$ the phase-space.

From the previous considerations, it is naturally to observe that, in order to uniquely determine the motion of the particle, we need to know, at least, its position at a fixed (initial) instant. That is we have to assign the following initial condition

$$y(t_0) = x_0$$

where $t_0 \in \mathbb{R}$ and $x_0 \in \mathbb{R}^k$ are fixed and such that $(t_0, x_0) \in \mathcal{D}$, the domain of f. Hence, we are assigning the value of the solution y at a fixed instant t_0. Then we have the following initial value first order system, more frequently called Cauchy problem

$$\begin{cases} y' = f(t, y) \\ y(t_0) = x_0. \end{cases}$$

Solving such a problem means to find a solution/trajectory which "passes" through x_0 at time t_0.

We end this subsection by applying all the definitions here given, to the equations of the examples of the previous subsection.

Equation (1.1) is: scalar, first-order, autonomous, linear homogeneous, in normal form.

Equation (1.2) is: scalar, second-order, autonomous, linear nonhomogeneous, in normal form.

Equation (1.3) is: scalar, first-order, autonomous, nonlinear, in normal form.

Equation (1.4) is: scalar, first-order, nonautonomous, nonlinear, in normal form.

Equation (1.5) is: a system of two first-order scalar equations, autonomous, nonlinear, in normal form.

Equation (1.6) is: scalar, second-order, autonomous, nonlinear, in normal form

Equation (1.7) is: scalar, first-order, autonomous, nonlinear, in non-normal form

1.3 Solving by hands and the need of a general theory

Let us consider the first order homogeneous nonautonomous linear equation

$$y'(t) = c(t)y(t), \quad t \in \mathbb{R}, \tag{1.13}$$

where $c : \mathbb{R} \to \mathbb{R}$ is a continuous function. Let $C : \mathbb{R} \to \mathbb{R}$ be a primitive of c. It is easy to see that, for every constant $k \in \mathbb{R}$, the function

$$y(t) = ke^{C(t)} \tag{1.14}$$

is a solution of (1.13).

The question is: are the functions of the form (1.14) all the solutions of (1.13)?

To answer the question, we are going to "work by hands" directly on the equation and then to get information on the solutions. First of all, we see that the null function $y \equiv 0$ is solution and it is of the form (1.14) with $k = 0$. Now, let y be a solution and let us suppose that there exists $t_0 \in \mathbb{R}$ such that $y(t_0) > 0$ (the case $y(t_0) < 0$ is similarly treated). Let $]a, b[\subseteq \mathbb{R}$ be the maximal interval such that

$$t_0 \in]a,b[\text{ and } y(t) > 0 \ \forall\, t \in]a,b[.$$

Starting from the equation (1.13) we then get

$$\frac{y'(t)}{y(t)} = c(t) \ \ \forall\, t \in]a,b[,$$

from which

$$\int_{t_0}^{t} \frac{y'(s)}{y(s)} ds = \int_{t_0}^{t} c(s) ds \ \ \forall\, t \in]a,b[.$$

Integrating and passing to the exponential, we finally get

$$y(t) = \left(y(t_0)e^{-C(t_0)}\right) e^{C(t)} \ \ \forall\, t \in]a,b[.$$

Hence, in the interval $]a,b[$, the solution y is of the form (1.14), with

$$k = \left(y(t_0)e^{-C(t_0)}\right) > 0. \tag{1.15}$$

Now, we observe that

$$a \in \mathbb{R} \implies y(a) = 0 \implies ke^{C(a)} = 0 \implies k = 0 \text{ contradiction!},$$

and similarly for b. Hence we must have $a = -\infty$ and $b = +\infty$, and we conclude that the solutions of (1.13) are exactly all the functions of the form (1.14)[23].

Now, the question is: which further conditions on the solution should we request, in order to uniquely fix the value of the constant k?

The answer is suggested by (1.15): we have to impose the value of the solution at a fixed time t_0, that is we have to consider the Cauchy problem

$$\begin{cases} y'(t) = c(t)y(t), \ t \in \mathbb{R}, \\ y(t_0) = y_0, \end{cases} \tag{1.16}$$

where $y_0 \in \mathbb{R}$ is the imposed value to the solution at $t = t_0$. It is now immediate to see that there is only one solution of the Cauchy problem (1.16): indeed we know that the solution is necessarily of the form (1.14) for some $k \in \mathbb{R}$, hence we get

$$y(t_0) = y_0 \implies ke^{C(t_0)} = y_0 \implies k = y_0 e^{-C(t_0)}.$$

Hence we have a unique solution to the problem (1.16), that is there exists a unique function y which solves (1.13) and, at the time t_0, passes through y_0. Such a function is[24]

[23] In particular, if $c(t) \equiv \tilde{c}$ is a constant, then the solutions are exactly all the functions of the form $y(t) = ke^{\tilde{c}t}$

[24] Again, if $c(\cdot)$ is the constant $\tilde{c}$, then we have $y(t) = y_0 e^{\tilde{c}(t-t_0)}$.

$$y(t) = \left(y_0 e^{-C(t_0)}\right) e^{C(t)} = y_0 e^{C(t)-C(t_0)}.$$

Let us summarize what we have discovered about the solutions of the equation (1.13) and of the Cauchy problem (1.16):

1) the solutions of (1.13) are the functions of the form $y_k(t) = ke^{C(t)}$ and they are defined (and solution) for all the times $t \in \mathbb{R}$;

2) for every $t_0, y_0 \in \mathbb{R}$ fixed, the solution of (1.16) is unique, and it is the function $y(t) = y_0 e^{C(t)-C(t_0)}$ (in particular, if $y_0 = 0$ then the unique solution is the null function $y \equiv 0$).

From 1) and 2) we can also get the following consideration

3) the general integral of (1.13), i.e. the set of all solutions, is a one-parameter family of functions

$$\tilde{I} = \{y_k : \mathbb{R} \to \mathbb{R} | k \in \mathbb{R}\},$$

and the correspondence $k \mapsto y_k$ between $\tilde{I}$ and $\mathbb{R}$ is a bijection. In particular, it is an injection because $k_1 \neq k_2$ implies $y_{k_1} \neq y_{k_2}$, since for instance they are different on t_0.

Now, we consider the following first order nonhomogeneous nonautonomous linear equation

$$y' = c(t)y + g(t), \tag{1.17}$$

where the function $c(\cdot)$ is as before and $g : \mathbb{R} \to \mathbb{R}$ is also continuous. Inspired by (1.14) in the previous case, we look for solutions of the form

$$y(t) = \alpha(t)e^{C(t)}, \tag{1.18}$$

where $\alpha : \mathbb{R} \to \mathbb{R}$ is a continuously derivable function. Differentiating and inserting in (1.17), we get

$$\alpha'(t) = g(t)e^{-C(t)}$$

and hence we must have

$$\alpha(t) = \int_{t_0}^{t} e^{-C(s)} g(s) ds + k,$$

where $t_0 \in \mathbb{R}$ is a fixed instant and $k \in \mathbb{R}$ is the integration constant. Hence, once t_0 is fixed, we have that, for all $k \in \mathbb{R}$, the function

$$y(t) = e^{C(t)} \left(k + \int_{t_0}^{t} e^{-C(s)} g(s) ds \right), \tag{1.19}$$

which is of the form (1.18), is a solution of (1.17).

Again, the question is: are the functions of the form (1.18) all the solutions of (1.17)?

Here, arguing as in the previous case is not completely immediate, due to presence of the term $g(t)$. However, using the already obtained existence and uniqueness results for the Cauchy problem (1.16), we immediate obtain a similar results for

$$\begin{cases} y'(t) = c(t)y(t) + g(t), \quad t \in \mathbb{R}, \\ y(t_0) = y_0, \end{cases} \tag{1.20}$$

where $t_0, y_0 \in \mathbb{R}$ are fixed values. Indeed, we easily get $k \in \mathbb{R}$ such that the function in (1.19) is a solution of (1.20):

$$y(t_0) = y_0 \implies k = y_0 e^{-C(t_0)}.$$

Now, using the linearity of (1.17), we have that, if $y(\cdot)$ and $z(\cdot)$ are any two solutions of (1.20), then the difference function $\psi = y - z$ is solution of (1.16) with condition $\psi(t_0) = 0$. But we already know that such a problem has a unique solution $\psi \equiv 0$. Hence we certainly have $y = z$, that is (1.20) has a unique solution, which of course is

$$y(t) = e^{C(t)} \left(y_0 e^{-C(t_0)} + \int_{t_0}^{t} e^{-C(s)} g(s) ds \right). \tag{1.21}$$

From such uniqueness result for (1.20), we can answer to the question whether all the solution of (1.17) are of the form (1.18). The answer is of course positive since, given any solution $y(\cdot)$ of (1.17) and denoted by y_0 its value in t_0, then such a function solves (1.20) and hence it is the function (1.21) which is of the form (1.18). Moreover, also in this case we get that the general integral of (1.17) is a one-parameter family of functions, one per every value of $k \in \mathbb{R}$.

What have we learned from the study of (1.17) and (1.20)? We have learned that, even if it is not obvious how to answer to our questions[25] via a direct hand-management of the equation, however we get a satisfactory answer using the already obtained uniqueness result for (1.16). In this, we are certainly helped by the fact that the equation is linear. Thus obtaining existence and uniqueness results for ordinary differential equations seems very important, even before making direct calculations for searching solutions. And what happens if the equation is not linear? Making the difference of two solutions is certainly not helpful. Hence, we are still more lead to think that a general and abstract theory concerning existence, uniqueness, comparison etc. of solutions is certainly useful and important. This is the subject of the next sections.

1.4 Frequently asked questions

Before starting with the general study of the ordinary differential equations, we make a list of those questions which are natural and common to formulate when we face an ordinary differential equation or a Cauchy problem.

Such questions are

[25] Are all the solutions of the form (1.18)? and is the solution of (1.20) unique?

(i) Does a solution exist?

(ii) If it exists, is it unique?

(iii) How long does a solution last?[26]

(iv) How much a solution is sensible with respect to parameters and coefficients which are present in the equation?

(v) How to calculate a solution?

Questions (i), (ii), (iii) and (iv) are of *qualitative* type. The question (v) is of *quantitative* type.

Another question is

(vi) Are there some types of equations which are easier to study than others, both from a qualitative and a quantitative point of view?

Another qualitative question is

(vii) If a solution exists for all times, what can we say about its behavior as time goes to infinity?

In the next sections, we will give some satisfactory answers to these questions. In particular questions (i) and (ii) are treated in Section 2; question (iii) is treated in Section 2 and Section 5; question (iv) is treated in Section 6 and Section 7; question (v) is treated in Section 4 and in Section 7; question (vi) is treated in Section 3; question (vii) is treated in Section 7.

[26] For instance, if we a priori fix a time interval $[t_1, t_2]$, does the solution exist for all those times?