LINEAR DIFFERENTIAL EQUATIONS

L. Boutet de Monvel
Université Pierre et Marie Curie, France

Keywords: Derivatives, differential equations, partial differential equations, distributions, Cauchy-Kowalewsky theorem, heat equation, Laplace equation, Schrödinger equation, wave equation, Cauchy-Riemann equations.

Contents

1. Introduction
2. Linearity and Continuity
2.1. Continuity
2.2. Linearity
2.3. Perturbation Theory and Linearity
2.4. Axiomatically Linear Equations
2.4.1. Fields: Maxwell Equations
2.4.2. Densities on Phase Space in Classical Physics
2.4.3. Quantum Mechanics and Schrödinger Equation
3. Examples
3.1. Ordinary Differential Equations
3.2. The Laplace Equation
3.3. The Wave Equation
3.4. The Heat Equation and Schrödinger Equation
3.5. Equations of Complex Analysis
3.5.1. The Cauchy-Riemann Equation
3.5.2. The Hans Lewy Equation
3.5.3. The Mizohata Equation
4. Methods
4.1. Well posed Problems
4.1.1. Initial Value Problem, Cauchy-Kowalewsky Theorem
4.1.2. Other Boundary Conditions
4.2. Distributions
4.2.1. Distributions
4.2.2. Weak Solutions
4.2.3. Elementary Solutions
4.3. Fourier Analysis
4.3.1. Fourier Transform
4.3.2. Equations with Constant Coefficients
4.3.3. Asymptotic Analysis, Microanalysis
Glossary
Bibliography
Biographical Sketch

Summary

Differential equations, as a part of differential calculus, were invented in the 17th

century by Leibniz and Newton; Newton illustrated the theory in a completely outstanding manner in his theory of gravitation and of the movement of planets. The theory of differential equations gives a means to reconstruct global objects, such as geometrical figures or movements, from their infinitesimal properties, i.e. their behavior in infinitely small regions of space or space-time (the position of the tangents, or relations between position, speed and acceleration). It is all-important in physics where one expects that phenomena propagate step by step and there is no « long distance » action (quantum theory obliges to somewhat revise this point of view).

This chapter deals with linear differential equations, i.e. of the form $df/dt-A(f)=g$, where the unknown function f and g are vector functions, and A depends linearly on f ($A(f+g)=A(f)+A(g)$). Linear equations form an important category because they appear systematically in error or perturbation computations (Section 2). They also appear in evolution processes which are axiomatically linear, i.e. where linearity enters in the definition, in particular in quantum theory. So they are a very important aspect of all evolution processes.

We have attempted to describe some of the most typical and important examples (Section 3), and the most performant methods that have been invented to deal with them (Section 4). The most important examples go back at least to Euler ; they describe the equilibrium state for an electric potential (Laplace equation), the propagation of light waves, or the diffusion of heat. We have added the more recent example of Hans Lewy, which belongs to complex geometry; the Hans Lewy equation cannot be solved, even locally, so it is a very bad model for physics; but it is very useful in complex analysis, and in some sense a generic model for « equations with non constant coefficients ». Among the most important tool used for the theory are the Fourier transformation and distribution theory. Fourier transformation describes the decomposition of a function or a movement as a superposition of plane waves (this is precisely what an oscilloscope does). The theory of distributions was developed by Sobolev and Schwartz ; its usefulness comes from the fact that for many differential problems concerning function of several variables, it is unavoidable to introduce functions which are less regular than differentiable function ; the theory of distributions gives a very practical and universal manner of dealing with such « irregular » objects ; it is essentially a « linear theory », and often more delicate to use in non-linear problems.

1. Introduction

A linear partial differential equation on $\mathbb{R}^n$ (n the number of variables) is an equation of the form

$$P(x,\partial) f = g ,$$

where $P = P(x,\partial) = \sum a_\alpha(x)\partial^\alpha$ is a linear differential operator, taking a differentiable function f into the linear combination of its derivatives:

$$\sum a_\alpha(x)\partial^\alpha f .$$

In this notation α is a multi-index i.e. a sequence of n integers $\alpha = (\alpha_1, \dots, \alpha_n)$ and the corresponding derivative is

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} \qquad (|\alpha| = \alpha_1 + \dots \alpha_n) \cdot$$

One also considers systems of such equations, where f (and the right hand side g) are vector valued functions and $P = (P_{ij})$ is a matrix of differential operators.

Linear equations appear systematically in error computations, or perturbation computations; they also appear in evolution processes which are axiomatically linear, i.e. where linearity enters in the definition, e.g. in field theories. So they are a very important aspect of all evolution processes.

For linear partial differential equations, a general theory was developed (although as yet certainly not complete). In many cases general rules of behavior, or for computing solutions exist, much more than for non linear partial differential equations, for which important mathematical (like functional analysis) were invented and used, but which consists mostly of a smaller number of fundamental systems of equations (like the Navier-Stokes equations describing the flow of a fluid) which model important physical phenomena and whose analysis also uses physical intuition.

2. Linearity and Continuity

2.1. Continuity

In our world, the state of a physical system is usually defined by a finite or infinite collection of real numbers (furnished by measures). These completely describe the state of the system if

1) The result of any other reasonable numerical measure one could perform on the system are determined by these numbers - i.e. in mathematical language, is a function of these, and

2) The future states of the system, i.e. the values of these numbers at future times $t \geq t_0$, are completely determined by these numbers at an initial time $t = t_0$. (In quantum physics things are a little different; one must use complex numbers, and the states of a system can no longer be thought of as a set or a manifold, where coordinates (measures) are real numbers; see below.) These numbers may satisfy some relations; mathematically one thinks of the set of all possible states as a manifold (possibly with singularities). For instance if the set of measures contains twice the same identical one, the corresponding numbers must of course be the same. There are many other possible ways of choosing measures that will have correlations, and in fact it will usually be impossible to give a complete description by a set of measures with no correlations at all. A simple example is a system whose states are points on a sphere, as on the surface of our planet Earth: a state can be

determined by 3 measures (height, breadth, depth) x, y, z satisfying one quadratic relation $x^2 + y^2 + z^2 = 1$. It is also determined by two real numbers (angles): the latitude and the longitude, but note however that the longitude is not well defined at the north and south poles.

Above we mentioned only "reasonable" measures. The reason is that our measures are never exact, there is always some error or imprecision (hopefully quite small) - this means that we can only make sense of measures that are continuous functions of the coordinates (the smaller the error on the coordinates, the smaller the error on the measure). Many usual functions are continuous. The elementary function $E(x) =$ integral part of x (largest integral number contained in x, for x a real number) is not continuous, and a computer is not really able to compute it: for instance the computer is not usually able to distinguish between the two numbers $x = 1 - \varepsilon$ and $y = 1 + \varepsilon$ if ε is a very small number (e.g. $\varepsilon = 10^{-43}$ = 1 over (1 followed by 43 zeros)), especially if these numbers are the result of an experiment and not given by a sure theoretical argument. Unless it is told to do otherwise, the computer will round up numbers and find $E(x) = E(y) = 1$; this means a huge relative error of 10^{43} (by comparison recall that the size of the known universe is about 10^{20} m).

2.2. Linearity

Measures provide real numbers. Real numbers can be added, and also multiplied (dilated). Objects or quantities for which addition and dilations are defined are usually called "vectors" and form a vector space; there is also a notion of complex vector spaces where dilations by complex numbers are allowed. This is not a complete mathematical definition; mathematicians also use, in particular in number theory or group theory, additive objects which have not much to do with vectors.

A real function of real numbers $f(x_1,\ldots,x_n)$ is linear if it takes sums into sums:

$$f(x_1 + y_1,\ldots,x_n + y_n) = f(x_1,\ldots,x_n) + f(y_1,\ldots,y_n)$$

(as limiting case f also takes a dilation to the same dilation (if it is continuous):

$$f(\lambda x_1,\ldots,\lambda x_n) = \lambda f(x_1,\ldots,x_n)$$

if λ is any real number)

Almost equivalently f is linear if it takes straight lines to straight lines or linear movements to linear movements (its "graph is straight")

Linear functions are simple to compute and to manipulate. The physical quantities we measure are not intrinsically linear; in fact the notion of linearity depends on the choice of basic coordinates (measures) one makes to describe a system, so it is not really well a priori defined for physical systems.

However one of the main points of the differential calculus which was developed during the 17th-18th centuries is the fact that usual functions or measurable physical phenomena are linear when restricted to infinitely small domains, or approximately linear when restricted to small domains (the smaller the domain, the better the relative linear approximation); in other words, error calculus is almost additive, the error produced by the superposition of two fluctuations in an experiment is the sum of the errors produced by each fluctuation separately (up to much smaller errors).

Functions which have this property of being approximately linear on small domains are called differentiable. The linear function which best approximates a differentiable function at a given point is called the tangent linear map, and its slope is the derivative at the given point. Most "usual" function given by simple explicit formulas have this property, e.g. the sine, exp, log function, and linear algebra appears systematically in error calculus concerning these functions.

Note however that Weierstrass showed that there are many continuous functions which are nowhere (or almost nowhere) differentiable - e.g. the function

$$y(x) = \sum 2^{-n} \sin 3^n x,$$

or the function which describes the shape of a coast: this is continuous, but its oscillations in small domains grow sharper and sharper. Many similar functions describing some kind of "fractal chaos" are in the same way continuous but not differentiable. Although they may be quite lovely to look at, they are difficult to handle and compute with quantitative precision - although small, the error on the result becomes comparatively very large when the increment of the variable is very small.

2.3. Perturbation Theory and Linearity

Anyway for usual "nice" functions the error calculus is always linear. Since the mathematical computation which to a differential equation or partial differential equation (and suitable boundary or initial data) assigns its solution looks nice (and at least in many important cases is nice), one also expects that if one makes a small perturbation of the equation or of the initial data, the resulting error on the solution will depend linearly on the errors on the data and one expects that it is governed by a system of linear differential equations. This is in fact true and not hard to prove in many good cases (although not always - e.g. there are equations without solutions, for which perturbation arguments do not make sense since there is nothing to begin with, see below). In any case, good or not, it is very easy, just using ordinary differential calculus, to write down the linear differential system that the error should satisfy.

For example if an evolution process is described by a differential equation depending on a parameter λ:

$$\frac{dx}{dt} = \Phi(t, x, \lambda) \qquad (1)$$

with initial condition at time $t = t_0$

$$x(t_0) = a(\lambda) \tag{2}$$

and $x_0(t)$ is a known solution for $\lambda = \lambda_0$, with initial value $x_0(t_0) = a(\lambda_0)$, then for close values $\lambda = \lambda_0 + \mu$, the solution $x = x(t,\lambda) = x_0 + u$ satisfies

$$\frac{dx}{dt} = \frac{dx_0}{dt} + \frac{du}{dt} = \Phi(t,x,\lambda) = \Phi(t,x_0,\lambda_0) + \nabla_x\Phi \cdot u + \nabla_\lambda\Phi \cdot \mu + \text{small error}$$

so the variation u satisfies, up to an infinitely small error, the linear equation:

$$\frac{du}{dt} = \nabla_x\Phi \cdot u + \nabla_\lambda\Phi \cdot \mu \tag{3}$$

with linear initial condition (with respect to μ):

$$u(t_0) = \nabla_\lambda a \cdot \mu \quad . \tag{4}$$

(the argument above, based on physical considerations, is intuitively convincing but it is still not a mathematical proof. In fact for differential equations (one variable), the proof is quite straightforward and taught in undergraduate courses. But although analogues for many good partial differential equations are true, for general partial differential equations the arguments above may be very hard to prove - and sometimes completely false.)

2.4. Axiomatically Linear Equations

Some theories are linear by their own nature, or axiomatically, so the differential equations that describe the behavior or the evolution of their objects must be linear.

2.4.1. Fields: Maxwell Equations

The description of several important physical phenomena uses fields and field theory. Fields are in the first place vector-functions (defined over time or space-time, or some piece of this, or some suitable space), but the point is that they are vector valued and make up a vector space. So equations for fields should be linear.

A striking example is the system of the Maxwell equations for the electro-magnetic field, in “condensed” form, and suitable units of time, length, electric charge, etc.

$$\nabla \cdot E = 4\pi\rho$$
$$\nabla \cdot B = 0$$
$$\nabla \times E + \frac{1}{c}\frac{dB}{dt} = 0$$
$$\nabla \times B - \frac{1}{c}\frac{dE}{dt} = \frac{4\pi}{c} j$$

In these equations, $E = (E_1, E_2, E_3)$, resp. $B = (B_1, B_2, B_3)$ is the electric, resp. magnetic field; c is the speed of light, ρ is the electric charge density, and j the electric current density, which is related to the charge by $\nabla j = d\rho/dt$.

For the vector field E with components (E_1, E_2, E_3), the usual notation $\nabla \cdot E$ (or $\operatorname{div} E$) means the scalar (one component) function

$$\nabla \cdot E = \frac{\partial E_1}{\partial x_1} + \frac{\partial E_2}{\partial x_2} + \frac{\partial E_3}{\partial x_3}$$

and $\nabla \times E = \operatorname{rot} E$ is the vector field with components

$$\nabla \times E = \begin{cases} \partial E_2/\partial x_3 - \partial E_3/\partial x_2 \\ \partial E_3/\partial x_1 - \partial E_1/\partial x_3 \\ \partial E_1/\partial x_2 - \partial E_2/\partial x_1 \end{cases}$$

the same notation is used for $B = (B_1, B_2, B_3)$.

We should note that these equations have a large group of symmetries: the Lorentz group, or the Poincaré group if one includes translations in space time; Lorentz transformations leave the equations invariant - not the solutions. These symmetries weave space and time together and were the starting point of relativity theory. The Maxwell equations are essentially the simplest system of first order linear which remain unchanged under transformations of the Poincaré group.

We add the following mathematical complement: let E be a vector space, and q a nondegenerate quadratic form (in the electromagnetic or relativistic setting $E = \mathbb{R}^4$, $q = t^2 - x^2 - y^2 - z^2$). Let Ω be the space of all differential forms on E (fields). A form of degree k can be written as $\sum a_{i_1 \ldots i_k}(x) dx_{i_1} \ldots dx_{i_k}$; the product of forms is defined and is anticommutative i.e. for odd forms $ba = -ab$.

The exterior derivation is the first order partial differential operator $d : \Omega \to \Omega$, which takes k-forms to the $k+1$-forms:

$$d \sum a_{i_1 \ldots i_k}(x) dx_{i_1} \ldots dx_{i_k} = \sum da_{i_1 \ldots i_k}(x) dx_{i_1} \ldots dx_{i_k}$$

with $df = \sum \frac{\partial f}{\partial x_i} dx_i$. The quadratic form d extends canonically to give a quadratic form or a scalar product on forms, and an infinitesimal volume element dv; the adjoint of d is the differential operator d^* such that

$$\int \langle d\omega \mid \phi \rangle = \int \langle \omega \mid d^* \phi \rangle .$$

One then defines the Dirac operator $D = d + d^*$ which takes even forms to odd forms, and is canonically associated to the quadratic form q. The system of Maxwell's equations is

$$DE = A \quad (E \in \Omega^{even}, A \in \Omega^{odd}) .$$

The electromagnetic field itself is described by 2-forms, and in the actual physical system certain components vanish - e.g. there are no magnetic charges.

-
-
-

Bibliography

Many encyclopaedias contain significant vulgarisation articles on differential or partial differential equations, in particular Encyclopaedia Britannica or Encyclopaedia Universalis.

The following are foundational books on PDE's, and any student or researcher on this subject must have read them.

Courant, R.; Hilbert, D. (1943) *Methoden der Mathematischen Physik, Vols. I, II*. Interscience Publishers, Inc., N.Y.

Gelfand, I.M.; Shilov, G.E. (1964, 1968, 1967) *Generalized functions, vol.1, 2, 3* (translated from Russian) Acad. Press New York & London.

Hadamard, J. Oeuvres. (1968) *Éditions du C.N.R.S.* Paris.

Hörmander, L. (1963) Linear Partial Differential Operators. Grundlehren der Math. Wiss. 116.

Hörmander, L. (1985)The analysis of linear partial differential operators vol. I, II, III, IV. Grundlehren der Math.Wiss. 256,257,274,275, Springer-Verlag.

John, F. (1955) *Plane waves and spherical means applied to partial differential equations*. Interscience, New York .

Schwartz, L. (1950,1951) *Théorie des Distributions, tome I et II*. Herman, Paris.

The following book of L. Schwartz, aimed at second or third year university students, is a remarkable and elementary introduction to P.D.E. theory

Schwartz, L. (1961) *Méthodes mathématiques de la physique*. Hermann, Paris,.

The following works were fundamental in the elaboration of the more recent point of view.

Guillemin V., Sternberg S. (1977) *Geometrical asymptotics. Amer. Math. Soc. Surveys 14,* Providence RI.

Leray J. (1957) Uniformisation de la solution du problème linéaire analytique de Cauchy près de la variété qui porte les données de Cauchy. *Bull. Soc. Math. France 85, 389-429.*

Maslov V.P. (1972) *Théorie des perturbations et Méthodes Asymptotiques*. Dunod , translated from Russian, Moscow University Editions, 1965).

Pseudo-differential operators and microlocal analysis (described in another chapter).

Duistermaat, J.J.; Hörmander, L. (1972) *Fourier integral operators II. Acta Math. 128, 183-269.*

Hörmander, L. (1965) *Pseudodifferential operators. Comm. Pure Appl. Math. 18, 501-517.*

Hörmander, L. (1972) *Fourier integral operators I. Acta Math. 127, 79-183.*

Kohn, J. J.; Nirenberg, L. (1965) An algebra of pseudo-differential operators. *Comm. Pure Appl. Math. 18 269–305.*

Analytic function theory.

Kashiwara, M.; Kawai, T.; Sato, M. (1973) *Microfunctions and pseudodifferential equations*, Lecture Notes 287, 265-524, Springer-Verlag.

Sato, M. Hyperfunctions and partial differential equations. (1969) *Proc. Int. Conf. on Funct. Anal. and Rel. topics, 91-94, Tokyo University Press, Tokyo.*

Non commutative geometry and non commutative analysis.

Connes, A. (1994) *Noncommutative geometry*. Academic Press, Inc., San Diego, CA.

Biographical Sketch

Louis Boutet de Monvel was born in 1941. He was student on the École Normale Supérieure in Paris, and was director of the E.N.S. mathematical center in 1979-86. He is presently professor in the university Pierre et Maris Curie (Paris 6), and in the Mathematical Institute, Jussieu. His research interests lie in the theory of linear partial differential equations and microlocal analysis, global and algebraic analysis. His main publications concern analytic pseudo-differential operators, boundary-value problems, Toeplitz operators and the Bergman kernel in complex analysis.