

PARTIAL DIFFERENTIAL EQUATIONS

SERGIU KLAINERMAN

1. BASIC DEFINITIONS AND EXAMPLES

To start with partial differential equations, just like ordinary differential or integral equations, are *functional equations*. That means that the unknown, or unknowns, we are trying to determine are functions. In the case of partial differential equations (PDE) these functions are to be determined from equations which involve, in addition to the usual operations of addition and multiplication, partial derivatives of the functions. The simplest example, which has already been described in section 1 of this compendium, is the Laplace equation in \mathbb{R}^3 ,

$$\Delta u = 0 \tag{1}$$

where $\Delta u = \frac{\partial^2}{\partial x^2} u + \frac{\partial^2}{\partial y^2} u + \frac{\partial^2}{\partial z^2} u$. The other two examples described in the section of fundamental mathematical definitions are the heat equation, with $k = 1$,

$$-\partial_t u + \Delta u = 0, \tag{2}$$

and the wave equation with $k = 1$,

$$-\partial_t^2 u + \Delta u = 0. \tag{3}$$

In these last two cases one is asked to find a function u , depending on the variables t, x, y, z , which verifies the corresponding equations. Observe that both (2) and (3) involve the symbol Δ which has the same meaning as in the first equation, that is $\Delta u = (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})u = \frac{\partial^2}{\partial x^2} u + \frac{\partial^2}{\partial y^2} u + \frac{\partial^2}{\partial z^2} u$. Both equations are called *evolution equations*, simply because they are supposed to describe the change relative to the time parameter t of a particular physical object. Observe that (1) can be interpreted as a particular case of both (3) and (2). Indeed solutions $u = u(t, x, y, z)$ of either (3) or (2) which are independent of t , i.e. $\partial_t u = 0$, verify (1).

A variation of (3), important in modern particle physics, is the Klein-Gordon equation, describing the free evolution, i.e. in the absence interactions, of a massive particle.

$$-\partial_t^2 u + \Delta u - m^2 u = 0. \tag{4}$$

Another basic equation of mathematical physics, which describes the time evolution of a quantum particle, is the Schrödinger equation,

$$i\partial_t u + k\Delta u = 0 \tag{5}$$

with u a function of the same variables (t, x, y, z) with values in the complex space \mathbb{C} and $k = \frac{\hbar}{2m} > 0$, where $\hbar > 0$ corresponds to the Planck constant and $m > 0$

the mass of the particle. As with our other two evolution equations, (2) and (3), above we simplify our discussion by taking $k = 1$.

Observe that all three PDE mentioned above satisfy the following simple property called the *principle of superposition*: If u_1, u_2 are solutions of an equation so is any linear combination of them $\lambda_1 u_1 + \lambda_2 u_2$ where λ_1 and λ_2 are arbitrary real numbers. Such equations are called linear. The following equation, called the minimal surfaces equation, is manifestly not linear. It refers to functions $u = u(x, y)$ which verify

$$\partial_x \left(\frac{\partial_x u}{(1 + |\partial_x u|^2 + |\partial_y u|^2)^{\frac{1}{2}}} \right) + \partial_y \left(\frac{\partial_y u}{(1 + |\partial_x u|^2 + |\partial_y u|^2)^{\frac{1}{2}}} \right) = 0. \quad (6)$$

Here ∂_x and ∂_y are short hand notations for the partial derivatives $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$.

The equations we have encountered so far can be written in the form $\mathcal{P}[u] = 0$, where \mathcal{P} is a *differential operator* applied to u . A differential operator is simply a rule which takes functions u , defined in \mathbb{R}^n or an open subset of it, into functions $\mathcal{P}[u]$ by performing the following operations:

- We can take partial derivatives $\partial_i u = \frac{\partial u}{\partial x^i}$ relative to the variables $x = (x^1, x^2, \dots, x^n)$ of \mathbb{R}^n . One allows also higher partial derivatives of u such as the mixed second partials $\partial_i \partial_j u = \frac{\partial^2 u}{\partial x^i \partial x^j}$ or $\partial_i^2 = \frac{\partial^2}{\partial x_i^2}$.
- The associated differential operators for (2) is $\mathcal{P} = -\partial_t + \Delta$ and that of (3) is $-\partial_t^2 + \Delta$
- Can add and multiply u and its partial derivatives between themselves as well as with given functions of the variables x . Composition with given functions may also appear.

In the case of the equation (1) the associated differential operator is $\mathcal{P} = \Delta = \partial_1^2 + \partial_2^2 + \partial_3^2 = \sum_{i,j=1}^3 e^{ij} \partial_i \partial_j$ where e^{ij} is the diagonal 3×3 matrix with entries $(1, 1, 1)$ corresponding to the euclidean scalar product of vectors X, Y in \mathbb{R}^3 ,

$$\langle X, Y \rangle = X_1 Y_1 + X_2 Y_2 + X_3 Y_3 = \sum_{i,j=1}^3 e^{ij} X_i Y_j. \quad (7)$$

The associated differential operators for (2), (3) and (5) are, resp. $\mathcal{P} = -\partial_t + \Delta$, $\mathcal{P} = -\partial_t^2 + \Delta$ and $\mathcal{P} = i\partial_t + \Delta$ with variables are $t, x^1, x^2, x^3 \in \mathbb{R}^{1+3}$. In the particular case of the wave equation (3) it pays to denote the variable t by x^0 . The wave operator can then be written in the form,

$$\square = -\partial_0^2 + \partial_1^2 + \partial_2^2 + \partial_3^2 = \sum_{\alpha,\beta=0}^3 m^{\alpha\beta} \partial_\alpha \partial_\beta \quad (8)$$

where $m^{\alpha\beta}$ is the diagonal 4×4 matrix with entries $(-1, 1, 1, 1)$, corresponding to the Minkowski scalar product in \mathbb{R}^{1+3} . This latter is defined, for 4 vectors $X = (X_0, X_1, X_2, X_3)$ and $Y = (Y_0, Y_1, Y_2, Y_3)$ by,

$$m(X, Y) = \sum_{\alpha,\beta=0}^3 m^{\alpha\beta} X_\alpha Y_\beta = -X_0 Y_0 + X_1 Y_1 + X_2 Y_2 + X_3 Y_3 \quad (9)$$

The differential operator \square is called D'Alembertian after the name of the French mathematician who has first introduced it in connection to the equation of a vibrating string.

Observe that the differential operators associated to the equations (1)–(5) are all linear i.e.

$$\mathcal{P}[\lambda u + \mu v] = \lambda \mathcal{P}[u] + \mu \mathcal{P}[v],$$

for any functions u, v and real numbers λ, μ . The following is another simple example of a linear differential operator

$$\mathcal{P}[u] = a_1(x)\partial_1 u + a_2(x)\partial_2 u \quad (10)$$

where $x = (x_1, x_2)$ and a_1, a_2 are given functions of x . They are called the *coefficients* of the linear operator. An equation of the form $\mathcal{P}[u] = f$, corresponding to a linear differential operator \mathcal{P} and a given function $f = f(x)$, is called *linear* even though, for $f \neq 0$, the principle of superposition of solutions does not hold.

In the case of the equation (6) the differential operator \mathcal{P} can be written, relative to the variables x^1 and x^2 , in the form,

$$\mathcal{P}[u] = \sum_{i=1}^2 \partial_i \left(\frac{1}{(1 + |\partial u|^2)^{\frac{1}{2}}} \partial_i u \right),$$

where $|\partial u|^2 = (\partial_1 u)^2 + (\partial_2 u)^2$. Clearly $\mathcal{P}[u]$ is not linear in this case. We call it a nonlinear operator; the corresponding equation (6) is said to be a *nonlinear equation*. An important property of both linear and nonlinear differential operators is *locality*. This means that whenever we apply \mathcal{P} to a function u , which vanishes in some open set D , the resulting function $\mathcal{P}[u]$ also vanishes in D .

Observe also that our equations (1)–(5) are also translation invariant. This means, in the case (1) for example, that whenever the function $u = u(x)$ is a solution so is the function $u_c(x) := u(T_c x)$ where T_c is the translation $T_c(x) = x + c$. On the other hand the equation $\mathcal{P}[u] = 0$, corresponding to the operator \mathcal{P} defined by (10) is not, unless the coefficients a_1, a_2 are constant. Clearly the set of invertible transformations $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ which map any solution $u = u(x)$, of $\mathcal{P}[u] = 0$, to another solution $u_T(x) := u(Tx)$ form a group, called the invariance group of the equation. The Laplace equation (1) is invariant not only with respect to translations but also rotations, i.e. linear transformations $O : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ which preserve the euclidean scalar product (7), i.e. $\langle OX, OY \rangle = \langle X, Y \rangle$ for all vectors $X, Y \in \mathbb{R}^3$. Similarly the wave equation (3) and Klein-Gordon equation (4) are invariant under Lorentz transformations, i.e. linear transformations $L : \mathbb{R}^{1+3} \rightarrow \mathbb{R}^{1+3}$ which preserve the Minkowski scalar product (9), i.e. $m(LX, LY) = m(X, Y)$. Our other evolution equations (2) and (5) are clearly invariant under rotations of the space variables $x = (x^1, x^2, x^3) \in \mathbb{R}^3$, keeping t fixed. They are also Galilean invariant, which means, in the particular case of the Schrödinger equation (5), that whenever $u = u(t, x)$ is a solution so is $u_v(t, x) = e^{i(x \cdot v)} e^{it|v|^2}(t, x - vt)$ for any vector $v \in \mathbb{R}^3$.

So far we have tacitly assumed that our equations take place in the whole space (\mathbb{R}^3 for the first, \mathbb{R}^4 for the second, third and fourth and \mathbb{R}^2 for the last example). In reality one is often restricted to a domain of the corresponding space. Thus,

for example, the equation (1) is usually studied on a bounded open domain of \mathbb{R}^3 subject to a specified *boundary condition*. Here is a typical example.

Example. The Dirichlet problem on an open domain of $D \subset \mathbb{R}^3$ consists of finding a continuous functions u defined on the closure \bar{D} of D , twice continuously differentiable in D , such that $\Delta u = 0$ in D and the restriction of u to ∂D , the boundary of D , is prescribed to be a continuous function u_0 . More precisely we require that,

$$u|_{\partial D} = u_0 \quad (11)$$

One can impose the same boundary condition for solutions of (6), with D a bounded open domain of \mathbb{R}^2 . A solution $u = u(x, y)$ of (6) in D , verifying the boundary condition (11), solves the Plateau problem of finding minimal surfaces in \mathbb{R}^3 which pass through a given curve. One can show that the surface given by the graph $\Gamma_u = \{(x, y, u(x, y)) / (x, y) \in D \subset \mathbb{R}^2\}$ has minimum area among all other graph surfaces Γ_v verifying the same boundary condition, $v|_{\partial D} = u_0$.

Natural boundary conditions can also be imposed for the evolution equations (2)–(5). The simplest one is to prescribe the values of u on the hyperplane $t = 0$. In the case of the heat and Schrödinger equation we set,

$$u|_{t=0} = u_0$$

while in the case of the wave equation we impose two conditions

$$u|_{t=0} = u_0 \text{ and } \partial_t u|_{t=0} = u_1 \quad (12)$$

where u_0, u_1 are functions of the coordinates (x, y, z) , called *initial conditions*. To solve the *initial value problem* in both cases means to find solutions of the equations for $t > 0$ which verify the corresponding initial conditions at $t = 0$. In addition one may restrict the variables (x, y, z) to an open domain of $D \subset \mathbb{R}^3$. More to the point one may try to solve a boundary value problem in a domain $[0, \infty) \times D$ with a boundary condition, such as (11), on $[0, \infty) \times \partial D$ and an initial condition at $t = 0$.

The choice of boundary condition and initial conditions, for a given PDE, is very important. Finding which are the good boundary and initial conditions is an important aspect of the general theory of PDE which we shall address in section 2. For equations of physical interest these appear naturally from the context in which they are derived. For example, in the case of a vibrating string, which is described by solutions of the one dimensional wave equation $\partial_t^2 u - \partial_x^2 u = 0$ in the domain $(a, b) \times \mathbb{R}$, the initial conditions $u = u_0, \partial_t u = u_1$ at $t = t_0$, amount to specifying the original position and velocity of the string. On the other hand the boundary condition $u(a) = u(b) = 0$ simply mean that the two ends of the of the string are fixed.

So far we have only considered equations in *one unknown*. In reality many of the equations of interest appear as systems of partial differential equations. The following important example, known as the Cauchy-Riemann equations contains two unknown functions $u_1 = u_1(x^1, x^2), u_2 = u_2(x^1, x^2)$ which verify

$$\partial_1 u_2 - \partial_2 u_1 = 0, \quad \partial_1 u_1 + \partial_2 u_2 = 0 \quad (13)$$

It was first observed by Cauchy that $u = u_1 + iu_2$, as a function of $z = x^1 + ix^2$, is a complex analytic function if and only if (13) is satisfied. Equation (13) can also be written in the form $\mathcal{P}[u] = 0$ by introducing $u = (u_1, u_2)$ as a column vector and $\mathcal{P}[u]$ the differential operator,

$$\mathcal{P}[u] = \begin{pmatrix} -\partial_2 & \partial_1 \\ \partial_1 & \partial_2 \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

The system of equations (13) contains two equations and two unknowns. This is the standard situation of a *determined system*. A system is called *over-determined* if it contains more equations than unknowns and *underdetermined* if it contains fewer equations than unknowns. Observe that (13) is a linear system. Observe also that the operator \mathcal{P} has the following remarkable property.

$$\mathcal{P}^2[u] = \mathcal{P}[\mathcal{P}[u]] = \begin{pmatrix} \Delta u_1 \\ \Delta u_2 \end{pmatrix}$$

In other words $\mathcal{P}^2 = \Delta \cdot I$, with I the identity operator $I[u] = u$, and therefore \mathcal{P} can be viewed as a square root of Δ . One can define a similar type of square root for the D'Alembertian \square . To achieve this we need 4×4 complex matrices $\gamma^0, \gamma^1, \gamma^3, \gamma^4$ which satisfy the property

$$\gamma^\alpha \gamma^\beta + \gamma^\beta \gamma^\alpha = -2m^{\alpha\beta} I \quad (14)$$

with I the unit 4×4 matrix and $m^{\alpha\beta}$ as in (8). Using the γ matrices we can introduce the Dirac operator acting on $u = (u_1, u_2, u_3, u_4)$ defined from \mathbb{R}^{1+3} with values in \mathbb{C}^4 by,

$$\mathcal{D}u = i\gamma^\alpha \partial_\alpha u \quad (15)$$

Using (14) we easily check that, $\mathcal{D}^2 u = \square u$. Thus the Dirac operator \mathcal{D} can be viewed as a square root of the D'Alembertian \square . The equation,

$$\mathcal{D}u = ku \quad (16)$$

is called the Dirac equation associated to a free, massive, relativistic, particle such as the electron.

Partial differential equations are ubiquitous throughout Mathematics and Science. They provide the basic mathematical framework for some of the most important physical theories, such as Elasticity, Hydrodynamics, Electromagnetism, General Relativity and Non-relativistic Quantum Mechanics. The more modern relativistic quantum field theories lead, in principle, to equations in infinite number of unknowns, which lie beyond the scope of partial differential equations. Yet, even in that case, the basic equations preserve the locality property of PDE. Moreover the starting point of a quantum field theory is always a classical field theory, described by systems of PDE's. This is the case, for example, of the Standard Model of weak and strong interactions, based on a Yang-Mills-Higgs field theory. If we also include the ordinary differential equations of Classical Mechanics, which can be viewed as one dimensional PDE, we see that, essentially, all of Physics is described by differential equations. As examples of partial differential equations underlining some of our most basic physical theories we refer to the articles of the compendium in which the Maxwell, Yang-Mills, Einstein, Euler and Navier Stokes equations are introduced.

Partial differential equations have also been found to play a fundamental role in areas of mathematics, which are considered pure, such as Complex Analysis, Differential Geometry, Topology and Algebraic Geometry. Complex Analysis, for example, which studies the properties of holomorphic functions, can be regarded as the study of solutions to the Cauchy-Riemann equations (13) in a domain of \mathbb{R}^2 . Hodge theory, based on studying the space of solutions to a class of linear systems of partial differential equations on manifolds which generalize the Cauchy-Riemann equations, plays a fundamental role in topology and algebraic geometry. The Atiyah-Singer index theorem is formulated in terms of a special classes of linear PDE on manifolds, related to the euclidean¹ version of the Dirac operator (15). Important geometric problems can be reduced to finding solutions to specific partial differential equations, typically nonlinear. We have already seen such an example in the case of the Plateau problem of finding surfaces of minimal total area which pass through a given curve. The well known uniformization theorem provides another excellent example. To state it we need to recall the definition of a compact Riemann surface S . This is a 2-dimensional, compact manifold endowed with a smooth, positive definite metric g . The Gauss curvature $K = K(g)$ is an important invariant of the surface which can be calculated explicitly at every point $p \in S$ in terms of the components g_{ab} relative to a local system of coordinates $x = (x^1, x^2)$ near p . The calculation involves first and second partial derivatives of the components g_{ab} relative to x^1 and x^2 . The remarkable fact is that the final value of K does not depend on the particular system of coordinates in which one makes the calculation. Moreover in the particular case when S is the standard sphere in \mathbb{R}^3 , given by the equation $|x|^2 = a^2$, the Gauss curvature is equal to the expected value, corresponding to our intuition of curvature, that is $K = a^{-2}$. Another remarkable property of the Gauss curvature is that its total integral along S does not depend on the metric g but only on the topological properties of S . More precisely, according to the Gauss-Bonnet formula, we have

$$\chi(S) = (2\pi)^{-1} \int_S K da_g,$$

with da_g denoting the area element of the metric g . In coordinates x^1, x^2 we have $da_g = \sqrt{|g|} dx^1 dx^2$ with $|g|$ the determinant of the matrix $(g_{ab})_{a,b=1,2}$. The number $\chi(S)$ is one of the integers $2, 0, -2, \dots -2k \dots$, called the Euler characteristic of S , and has simple topological interpretation. Thus any surface which can be continuously deformed to the standard sphere has $\chi(S) = 2$ while any surface which can be continuously deformed to a torus has $\chi(S) = 0$. We can now state the uniformization theorem:

Theorem 1. *Let S be a 2-dimensional, compact, Riemann surface with metric g , Gauss curvature $K = K(g)$ and Euler characteristic $\chi(S)$. There exists a conformal transformation of the metric g , i.e. $\tilde{g} = \Omega^2 g$, for some smooth non-vanishing function Ω , such that the Gauss curvature \tilde{K} of the new metric \tilde{g} is identical equal to 1, 0 or -1 according to whether $\chi(S) > 0$, $\chi(S) = 0$ or $\chi(S) < 0$.*

To prove this very important geometric result, which leads to the complete classification of all compact surfaces according to their Euler characteristic, we are led

¹This is the operator obtained when we change the minkowski metric m to the euclidean one e in (14).

to a nonlinear partial differential equation on S . Indeed assume that $\chi(S) = 2$ and therefore we want the Gauss curvature \tilde{K} of the metric $\tilde{g} = e^{2u}g$ to be exactly 1. It is easy to relate \tilde{K} , by a simple calculation, to the Gauss curvature K of the original metric g . This leads to the following equation in u ,

$$\Delta_S u + e^{2u} = K \quad (17)$$

where Δ_S , called the Laplace-Beltrami operator of S , is a straightforward adaptation of the Laplace operator, see (1), to the surface S . Thus the proof of the uniformization theorem reduces to solve equation (17), i.e. for a given surface S with Gauss curvature K , find a real valued function u which verifies (17).

We give below a precise definition of the operator Δ_S relative to a system of local coordinates $x = (x^1, x^2)$ on an open coordinate chart $D \subset S$. Denote by $G(x) = (g_{ab}(x))_{a,b=1,2}$ the 2×2 matrix whose entries are the components of our Riemannian metric on D . Let $G^{-1}(x)$ denote the matrix inverse to $G(x)$ and denote its components by $(g^{ab}(x))_{a,b=1,2}$. Thus, for all $x \in D$,

$$\sum_c g_{ac}(x)g^{cb}(x) = \delta_{ab}$$

with δ_{ab} the usual Kronecker symbol. We also set, as before, $|g(x)| = \det(G(x))$ and define,

$$\Delta_S u(x) = \frac{1}{\sqrt{|g(x)|}} \sum_{a,b=1,2} \partial_b(\sqrt{|g(x)|} g^{ab}(x) \partial_a u(x))$$

Typically we suppress the explicit dependence on x in the above formula. It is also very convenient to use Einstein's summation convention over repeated indices, and thus write,

$$\Delta_S u = \frac{1}{\sqrt{|g|}} \partial_b(\sqrt{|g|} g^{ab} \partial_a u) \quad (18)$$

As a third example we consider the Ricci flow equation on a compact n dimensional manifold M , which is described in one of the articles of the Compendium. In the particular case of three dimensions the equation has been recently used, decisively, to provide the first proof of Thurston's geometrization conjecture, including the well known Poincaré conjecture. The geometrization conjecture, described in the topology section of the Compendium, is the precise analogous, in three space dimensions, of the 2-dimensional uniformization theorem mentioned above. The Ricci flow is defined, in arbitrary local coordinates $x = (x^1, x^2, x^3)$ on M , by the equation:

$$\partial_t g_{ij} = R_{ij}(g) \quad (19)$$

Here $g_{ij} = g_{ij}(t)$ is a family of Riemannian metrics depending smoothly on the parameter t and $R_{ij}(g)$ denotes the Ricci curvature of the metric g_{ij} . This is simply a three dimensional generalization of the Gauss curvature we have encountered in the uniformization theorem. In a given system of coordinates $R_{ij}(g)$ can be calculated in terms of the metric coefficients g_{ij} and their first and second partial derivatives. Since both g_{ij} and R_{ij} are symmetric relative to $i, j = 1, 2, 3$ we can interpret (19) as a non-linear system of six equations with six unknowns. On a closer look it turns out that (19) is related to the heat equation (2). Indeed, by

a straightforward calculation relative to a particular system of coordinates $x = (x^1, x^2, x^3)$ called *harmonic*, it can be shown that the Ricci flow (19) takes the form

$$\partial_t g_{ij} - \Delta_g g_{ij} = N_{ij}(g, \partial g) \quad (20)$$

where each N_{ij} , $i, j = 1, 2, 3$, are functions of the components g_{ij} and their first partial derivatives with respect to the coordinates x and Δ_g is, again, a differential operator very similar to the Laplacean Δ in \mathbb{R}^3 , see (1). More precisely, if $G^{-1} = (g^{ab})_{a,b=1,2,3}$ denotes the matrix inverse to $G = (g_{ab})_{a,b=1,2,3}$ we can write, using the summation convention,

$$\Delta_g = g^{ab} \partial_a \partial_b = \sum_{a,b=1}^3 g^{ab} \partial_a \partial_b.$$

In a small neighborhood of a point $p \in M$ we can choose the harmonic coordinate x^a such that $g^{ab}(p) = \delta^{ab}$ with δ^{ab} denoting the usual Kronecker symbol. Thus, near p , Δ_g looks indeed like $\Delta = \delta^{ab} \partial_a \partial_b$.

The Ricci flow² allows one to deform an arbitrary Riemannian metric on M to a simple metric of constant sectional curvature. The idea is to start with a metric g and look for solutions $g(t)$ of (19) which verify the *initial condition* $g(0) = g$. One hopes that as $t \rightarrow \infty$ the metric $g(t)$ will converge to a metric of constant curvature. Intuitively one can see this working out the same way heat gets evenly distributed in space, as t increases, according to the heat equation (2). Indeed since (19) is similar to (2) we expect the variations in the curvature of $g(t)$ to become smaller and smaller as the metric evolves according to (19). The type of metric we get in the limit as $t \rightarrow \infty$ will allow us to determine the topological character of M . The flow, however, can *develop singularities* before we achieve that goal. To overcome this major technical difficulty one needs to make a detailed qualitative analysis of the behavior of solutions to (19), task which requires just about all the advances made in geometric PDE in the last hundred years.

As we have seen above the choice of harmonic coordinates allows us to write the Ricci flow as a system of *nonlinear heat equations* (20). This fact is quite typical to geometric equations. It is useful at this point to discuss another important example, that of the Einstein equations in vacuum. An introduction to this equation and short discussion of its importance in General Relativity can be found (see compendium article). Solutions to the Einstein vacuum equations are given by Ricci flat *spacetimes*, that is Lorentzian manifolds (M, g) with M a four dimensional manifold and g a Lorentz metric on it, for which the corresponding Ricci curvature vanishes identically. The Ricci curvature of a Lorentz metric can be defined in exactly the same way as in the Riemannian case. Thus relative to a coordinate system x^α , with $\alpha = 0, 1, 2, 3$, the Ricci curvature $R_{\alpha\beta}$ can be expressed in terms of the first and second partial derivatives of the metric coefficients $g_{\alpha\beta}$. As before, we denote by $g^{\alpha\beta}$ the components of the inverse metric. Moreover, by picking a specified system of coordinates, called *wave coordinates*³, we can express the

²In reality one needs to change the equation (19) slightly to make sure that the total volume of M , calculated with respect to the metric $g(t)$, stays constant.

³they are the exact analogue of the harmonic coordinates discussed above.

Einstein-vacuum equations,

$$R_{\alpha\beta} = 0, \tag{21}$$

in the form of a system of equations related to the wave equation (3), in the same way the Ricci flow system (20) was related to the heat equation (2). More precisely,

$$\square_g g_{\alpha\beta} = N_{\alpha\beta}(g, \partial g) \tag{22}$$

where, as in the case of the Ricci flow, the terms $N_{\alpha\beta}(g, \partial g)$ are expressions, which can be calculated explicitly, depending on the metric $g_{\alpha\beta}$, its inverse $g^{\alpha\beta}$ and the first derivatives of $g_{\alpha\beta}$ relative to the coordinates x^α . This is a system of 10 equations with respect to the ten unknown components of the metric $(g_{\alpha\beta})_{\alpha,\beta=0,1,2,3}$. The differential operator,

$$\square_g = \sum_{\mu,\nu} g^{\mu\nu} \partial_\mu \partial_\nu$$

appearing on the left hand side is very similar to the wave operator $\square = m^{\mu\nu} \partial_\mu \partial_\nu = -\partial_0^2 + \Delta$ which we have encountered before in (8). Indeed, in a neighborhood of a point $p \in M$ we can pick our wave coordinates x^α in such a way that $g^{\mu\nu}(p) = m^{\mu\nu}$. Thus, locally, \square_g looks like $\square = \square_m$ and we can thus interpret (22) as a nonlinear system of *wave equations*.

The two last examples illustrate the importance of choosing *good coordinates* for equations which are defined in terms of geometric quantities, such as the Ricci curvature. To solve such equations and find interesting properties of the solutions, it is often very important to pick up a well adapted system of coordinates. In the case of gauge field theories, such as Yang-Mills equations, the role of coordinates is replaced by *gauge transformations*.

Finally we need to note that PDE arise not only in Physics and Geometry but also in many fields of applied science. In engineering, for example, one often wants to impose auxiliary conditions on solutions of a PDE, corresponding to a part of a physical system which we can directly influence, such as the portion of the string of a violin in direct contact with the bow, in order to control their behavior, i.e. obtain a beautiful sound. The mathematical theory dealing with this issue is called *Control Theory*.

Often, when dealing with complex physical systems, when we cannot possible have complete information about the state of the system at any given time, one makes various randomness assumptions about various factors which influence it. This leads to a very important class of equations called stochastic differential equations. To give a simple example consider⁴ the $N \times N$ system of the ordinary differential equation,

$$\frac{dx}{dt} = f(x(t)) \tag{23}$$

Here f is a given function $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$. A solution $x(t)$ is a vector valued function $x : [0, \infty) \rightarrow \mathbb{R}^N$. Given an initial data $x(0) = x_0$ we can precisely determine the position $x(t)$ and velocity $\frac{dx}{dt}$ of the solution at any given time t . In applied

⁴For simplicity the reader can just take $N = 1$.

situations, because of various factors which are hard to take into account, the state of the solution may not be so neatly determined. It is thus reasonable to modify the equation to take into account random effects which influence the system. One then looks at an equation of the form,

$$\frac{dx}{dt} = f(x(t)) + B(x(t))\frac{dW}{dt}(t) \quad (24)$$

where $B(x)$ is a $N \times M$ dimensional matrix and $W(t)$ denotes the brownian motion in \mathbb{R}^M . Similar modifications, which take randomness into account, can be made for partial differential equations.

A particularly interesting example of a PDE, which is derived from a stochastic process, related to the price of stock options in finance, is the well known Black-Scholes equation. The real price of a stock option $u(s, t)$ at time t and value s , verifies the PDE,

$$\partial_t u + rs\partial_s u + \frac{\sigma^2}{2}s^2\partial_s^2 u - ru = 0, \quad s > 0, \quad t \in [0, T], \quad (25)$$

subject to the terminal condition at expiration time T , $u = \max(0, (s - p))$ and boundary condition $u(0, t) = 0, t \in [0, T]$. Here p is the strike price of the option.

Here is the plan for the rest of the article. In section 2 we describe some of the basic notions and achievements of the general theory of PDE. This is the only section which is somewhat abstract; the main point I want to make here is that unlike ordinary differential equations, for which a general theory is both possible and useful, a general, useful, theory for PDE is unfortunately not possible due to some important obstructions which I will try to describe. One is thus forced to discuss, in some generality, classes of equations such as *elliptic, parabolic, hyperbolic and dispersive*. In section 3, I will try to argue that, despite the obvious impossibility to develop a useful general theory which encompasses all, or most, of the important examples, there exists nevertheless an impressive, unifying, body of concepts and methods, for dealing with various basic equations, which gives PDE the aspect of a well defined area of mathematics. In section 4 we develop this further by trying to identify some common features in the derivation of the main equations we deal with in our subject. In chapter 5 we discuss one of the most fundamental problem in PDE, that of regularity or break-down of solutions. Finally in sections 6 and 7 I try to identify some of the main goals and open problems in PDE.

2. GENERAL EQUATIONS

We may be tempted to define PDE as the subject which is concerned with all partial differential equations. According to this view, the goal of the subject is to find a general theory of all, or very general classes of PDE's. As we shall argue below, this point of view is seriously flawed and very much out of fashion. It has nevertheless important merits which I hope to illustrate below. To see the full power of the general theory I need to, at least, write down general equations, yet I make sure to explain the main ideas in simplified cases. I consider equations, or systems of equations, in \mathbb{R}^d with respect to the variables $x = (x^1, x^2, \dots, x^d)$. We

denote by $\partial_i = \frac{\partial}{\partial x^i}$ the partial derivatives relative to the coordinate x^i and by $\partial^\alpha = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_d^{\alpha_d}$ the mixed partial derivatives corresponding to a multi-index $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{N}^d$. We denote by ∂^k the vector of all partial derivatives ∂^α with $|\alpha| = \alpha_1 + \dots + \alpha_d = k$. Finally we denote by $\Lambda^k u = (u, \partial u, \dots, \partial^k u)$ the set of all partial derivatives of order less or equal to k . In most interesting examples k is one or two.

Example. To make these notations more transparent consider the case of \mathbb{R}^2 and coordinates x^1, x^2 . For the multi-index $\alpha = (2, 0)$ we have $\partial^\alpha u = \frac{\partial}{\partial x^1} \frac{\partial}{\partial x^1} u = \partial_1^2 u$ while for $\alpha = (1, 1)$ we have $\partial^\alpha u = \frac{\partial}{\partial x^1} \frac{\partial}{\partial x^2} u = \partial_1 \partial_2 u$. Also

$$\partial^2 u = \left(\frac{\partial}{\partial x^1} \frac{\partial}{\partial x^1} u, \frac{\partial}{\partial x^1} \frac{\partial}{\partial x^2} u, \frac{\partial}{\partial x^2} \frac{\partial}{\partial x^2} u \right) = (\partial_1^2 u, \partial_1 \partial_2 u, \partial_2^2 u)$$

and $\Lambda^2 u = (u, \partial_1 u, \partial_2 u, \partial_1^2 u, \partial_1 \partial_2 u, \partial_2^2 u)$.

With this notation the Laplace operator in \mathbb{R}^d has the form $\Delta = \partial_1^2 + \partial_2^2 + \dots + \partial_d^2$ while the D'Alembertian in the Minkowski space \mathbb{R}^{d+1} has the form $\square = -\partial_t^2 + \partial_1^2 + \dots + \partial_d^2$. To make sense of an equation in which there appear partial derivatives of order up to k we need to work with functions which are k -time differentiable at every point. It is convenient to work with the class C^k of functions which are continuous and whose all partial derivatives $\partial^\alpha u$ of order $|\alpha| \leq k$ are continuous.

Definition. A general partial differential equation in \mathbb{R}^d of order k is of the form,

$$F(x, \Lambda^k u(x)) = 0, \tag{26}$$

where F is a specified function. We also consider $N \times N$ systems⁵ in which case F and u are column N -vectors. A function u of class C^k is said to be a classical⁶ solution (26) if it verifies the equation as all points x in a specified domain of \mathbb{R}^d .

Consider first the one dimensional situation $d = 1$ in which case (26) becomes an ordinary differential equation (ODE), or system of ODE. To simplify further take $k = 1$ and $N = 1$, that is the case of an ordinary differential equation of order $k = 1$. Then (26) is simply, $F(x, u(x), \partial_x u(x)) = 0$ where F is a given function of the three variables x, u and $p = \partial_x u$ such as, for example, $F(x, u, p) = x \cdot p + u^3 - \sin x$. To solve the equation (26) in this case is to find a function a C^1 function $u(x)$ such that

$$x \cdot \partial_x u(x) + u^3 = \sin x. \tag{27}$$

Now consider the case of a second order ODE, i.e. $d = N = 1$ and $k = 2$. Then (26) becomes, $F(x, u(x), \partial_x u(x), \partial_x^2 u(x)) = 0$, where F now depends on the four variables $x, u, p = \partial_x u, q = \partial_x^2 u$. As an example take $F = q^2 + V'(u)$, for some given function $V = V(u)$, in which case (26) becomes the nonlinear harmonic oscillator equation,

$$\partial_x^2 u(x) + V'(u(x)) = 0 \tag{28}$$

Passing to a system of ODE, with $d = 1, k = 1$ and $N = 2$ we will need a vector function $F = (F_1, F_2)$ with both F_1 and F_2 depending on the five variables

⁵That is *determined systems* of N equations for N unknowns.

⁶We call it classical to distinguish from generalized solutions to be discussed in the following sections.

$x, u_1, u_2, p_1 = \partial_x u_1, p_2 = \partial_x u_2$. Then (26) becomes,

$$\begin{aligned} F_1(x, u_1(x), u_2(x), \partial_x u_1(x), \partial_x u_2(x)) &= 0 \\ F_2(x, u_1(x), u_2(x), \partial_x u_1(x), \partial_x u_2(x)) &= 0 \end{aligned}$$

The case of PDE gets a bit more complicated because of the large number of variables involved in the definition of F . Thus for first order ($k = 1$) scalar equations ($N=1$) in two space dimensions ($d = 2$) we need functions F depending on the two spatial variables x^1, x^2 as well as $u, p_1 = \partial_1 u$ and $p_2 = \partial_2 u$. For a given function of five variables $F = F(x, u, p)$, a general first order PDE in two space dimensions takes the form,

$$F(x, u(x), \partial_1 u(x), \partial_2 u(x)) = 0. \quad (29)$$

As a particular example take $F = p_1^2 + p_2^2 - 1$. The corresponding equation is,

$$(\partial_1 u(x))^2 + (\partial_2 u(x))^2 = 1 \quad (30)$$

which plays an important role in geometric optics. A classical solution of the equation is a C^1 function $u = u(x^1, x^2)$ which verifies (30) at all points of a domain $D \subset \mathbb{R}^2$.

Remark 1. We have excluded from our definition over-determined (i.e. the number of equations exceeds that of unknowns) or underdetermined systems (i.e. the number of equations is less than that of unknowns) despite their obvious interest to Geometry and Physics. The Einstein vacuum equations (21), for example, look underdetermined at first glance. They become determined once we fix a particular coordinate condition, such as the *wave coordinate* condition alluded to in section 1. Gauge field theories, such as Yang-Mills, have a similar structure.

Remark 2. All higher order scalar equations or systems can in fact be re-expressed as first order systems, i.e. $k = 1$, by simply introducing all higher order derivatives of u as unknowns together with the obvious compatibility relations between partial derivatives. As an example consider equation (28) and set $v = \partial_x u$. We can then rewrite the equation as a first order system with $N = 2$, namely $\partial_x v + V'(u) = 0, \partial_x u - v = 0$.

An equation, or system, is called *quasi-linear* if it is linear with respect to the highest order derivatives. A quasilinear system of order one ($k = 1$) in \mathbb{R}^d can be written in the form,

$$\sum_{i=1}^d A^i(x, u(x)) \partial_i u = F(x, u(x)) \quad (31)$$

Here u and F are column N -vectors and the coefficients A^1, A^2, \dots, A^d are $N \times N$ matrix valued functions.

The minimal surface equation is an example of a second order ($k = 2$) quasilinear scalar equation ($N = 1$) in two space dimensions. Indeed, using the coordinates x^1, x^2 , instead of x, y , we can manipulate (6) with the help of Leibnitz formula and

rewrite in the form,

$$\sum_{i,j=1,2} h^{ij}(\partial u) \partial_i \partial_j u = 0, \quad (32)$$

with $h^{11}(\partial u) = 1 + (\partial_2 u)^2$, $h^{22}(\partial u) = 1 + (\partial_1 u)^2$, $h^{12}(\partial u) = h^{21}(\partial u) = -\partial_1 u \cdot \partial_2 u$, which is manifestly a second order quasi-linear equation.

In the particular case when the top order coefficients of a quasilinear equation, i.e. those corresponding to the highest order derivatives, depend only on the space variables $x \in \mathbb{R}^d$, the equation, or system, is called semi-linear. For example, equation (17) derived in connection to the uniformization theorem, is semi-linear.

A linear equation, or system, of order k can be written in the form,

$$\sum_{|\alpha| \leq k} A_\alpha(x) \partial^\alpha u(x) = F(x). \quad (33)$$

Observe that the differential operator on the left hand side is indeed linear in the sense discussed in our introduction. If in addition the coefficients A_α are constant in x , the system is called linear with constant coefficients. The five basic equations (1)–(5) discussed in the introduction are all linear with constant coefficients. Typically, these are the only equations which can be solved explicitly.

We thus have our first useful, indeed very useful, classification of PDE's into fully nonlinear, quasi-linear, semi-linear and linear. A fully nonlinear equation is nonlinear relative to the highest derivatives. The typical example is the Monge Ampere equation. For simplicity consider the case of functions of 2 variables $u(x^1, x^2)$ in \mathbb{R}^2 with hessian $\partial^2 u = (\partial_i \partial_j u)_{i,j=1,2}$. Clearly the determinant $\det(\partial^2 u) = (\partial_1^2 u) \cdot (\partial_2^2 u) - (\partial_1 \partial_2 u)^2$, is quadratic with respect to the second derivatives of u . Thus the Monge -Ampère equation,

$$\det(\partial^2 u) = f(x, u, \partial u), \quad (34)$$

with f a given function defined on $\mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2$, is fully nonlinear. This equation plays an important role in Geometry, in relation to the isometric embedding problem as well as to the problem of finding surfaces in \mathbb{R}^2 with prescribed Gauss curvature. A variant of the Monge Ampère equation, for complex valued functions, plays a central role in complex geometry in connection to Calabi -Yau manifolds. Calabi-Yau manifolds, on the other hand, are central mathematical objects in String Theory.

Remark. Most of the basic equations of Physics, such as the Einstein equations, are quasilinear. Fully nonlinear equations appear however in connection to the theory of characteristics of linear PDE, which we discuss at length below, or in geometry.

1. *First order scalar equations.* It turns out that scalar ($N = 1$) first order ($k = 1$) PDE in d space dimensions can be reduced to systems of first order ODE. As a simple illustration of this important fact consider the following equation in two space dimensions,

$$a^1(x^1, x^2) \partial_1 u(x^1, x^2) + a^2(x^1, x^2) \partial_2 u(x^1, x^2) = f(x^1, x^2) \quad (35)$$

where a^1, a^2, f are given real functions in the variables $x = (x^1, x^2) \in \mathbb{R}^2$. We associate to (35) the first order 2×2 system

$$\frac{dx^1}{ds}(s) = a^1(x^1(s), x^2(s)), \quad \frac{dx^2}{ds} = a^2(x^1(s), x^2(s)) \quad (36)$$

To simplify matters we assume $f = 0$. Observe that any solution $u = u(x^1, x^2)$ of (35), with $f = 0$, is constant along any solution $x(s) = (x^1(s), x^2(s))$, i.e. $\frac{d}{ds}u(x^1(s), x^2(s)) = 0$. Thus, in principle, the knowledge of solutions to (36), which are called *characteristic curves* for (35), allows us to find all solutions to (35). I say in principle because, in general, the nonlinear system (36) is not so easy to solve. Yet ODE are simpler to deal with and the fundamental theorem of ODE, which we will discuss later in this section, allows us to solve (36), at least locally for a small interval in s . The constancy of u along characteristic curves allows us to obtain, even when we cannot find explicit solutions, important qualitative information. For example, suppose that the coefficients a^1, a^2 are smooth (or real analytic) and that the initial data is smooth (or real analytic) everywhere on \mathcal{H} except at some point $x_0 \in \mathcal{H}$ where it is discontinuous. Then, clearly, the solution u remains smooth (or real analytic) at all points except along the characteristic curve Γ which initiates at x_0 , i.e. along the solution to (36) which verifies the initial condition $x(0) = x_0$. The discontinuity at x_0 propagates precisely along Γ . We see here the simplest manifestation of a general principle, which we shall state later, that singularities of solutions to PDE propagate along characteristics.

One can generalize equation (35) to allow the coefficients a_1, a_2 and f to depend not only on $x = (x^1, x^2)$ but also on u ,

$$a^1(x, u(x))\partial_1 u(x) + a^2(x, u(x))\partial_2 u(x) = f(x, u(x)) \quad (37)$$

The associated *characteristic system* becomes,

$$\frac{dx^1}{ds}(s) = a^1(x(s), u(s, x(s))), \quad \frac{dx^2}{ds} = a^2(x(s), u(s, x(s))) \quad (38)$$

As a special example of (37) consider the scalar equation in two space dimensions,

$$\partial_t u + u\partial_x u = 0, \quad u(0, x) = u_0(x) \quad (39)$$

called the Burger equation. Since $a^1 = 1, a^2 = u$ we can set $x^1(s) = s, x^2(s) = x(s)$ in (38) and thus derive its *characteristic equation* in the form,

$$\frac{dx}{ds}(s) = u(s, x(s)). \quad (40)$$

Observe that, for any given solution u of (39) and any characteristic curve $(s, x(s))$ we have $\frac{d}{ds}u(s, x(s)) = 0$. Thus, in principle, the knowledge of solutions to (40) would allow us to determine the solutions to (39). This, however, seems circular since u itself appears in (40). To see how this difficulty can be circumvented consider the initial value problem for (39), i.e. look for solutions u which verify $u(0, x) = u_0(x)$. Consider an associated characteristic curve $x(s)$ such that, initially, $x(0) = x_0$. Then, since u is constant along the curve, we must have $u(s, x(s)) = u_0(x_0)$. Hence, going back to (40), we infer that $\frac{dx}{ds} = u_0(x_0)$ and thus $x(s) = x_0 + su_0(x_0)$. We thus deduce that,

$$u(s, x_0 + su_0(x_0)) = u_0(x_0) \quad (41)$$

from which gives us, implicitly, the form of the solution u . We see once more, from (41), that if the initial data is smooth (or real analytic) everywhere except at a point x_0 , of the line $t = 0$, then the corresponding solution is also smooth (or real analytic) everywhere, in a small neighborhood V of x_0 , except along the characteristic curve which initiates at x_0 . The smallness of V is necessary here because new singularities can form in the large. Observe indeed that u has to be constant along the lines $x + su_0(x)$ whose slopes depend on $u_0(x)$. At a point when these lines cross, we would obtain different values of u which is impossible unless u becomes singular at that point. In fact one can show that the first derivative u_x becomes infinite at the first singular point, i.e. the singular point with the smallest value of $|t|$. This blow-up phenomenon occur for any smooth, non-constant, initial data u_0 .

Remark. There is an important difference between the linear equation (35) and quasi-linear equation (37). The characteristics of the first depend only on the coefficients $a^1(x), a^2(x)$ while the characteristics of the second depend, explicitly, on a particular solution u of the equation. In both cases, *singularities can only propagate along the characteristic curves of the equation*. For nonlinear equations, however, new singularities can form in the large, independent of the smoothness of the data.

The above procedure extends to fully nonlinear scalar equations in \mathbb{R}^d such as,

$$\partial_t u + H(x, \partial u) = 0, \quad u(0, x) = u_0(x) \quad (42)$$

with $H = H(x, p)$ a given function of the variables $x = (x^1, x^2, \dots, x^d)$ and $p = (p_1, p_2, \dots, p_d)$, called the *Hamiltonian* of the system, and $\partial u = (\partial_1 u, \partial_2 u, \dots, \partial_d u)$. We associate to (42) the ODE system, with $i = 1, 2, \dots, d$,

$$\frac{dx^i}{dt} = \frac{\partial}{\partial p_i} H(x(t), p(t)), \quad \frac{dp_i}{dt} = -\frac{\partial}{\partial x^i} H(x(t), p(t)). \quad (43)$$

The equation (42) is called a Hamilton-Jacobi equation while (43) is known as a Hamiltonian system of ODE. The relationship between them is a little more involved than in the previous cases discussed above. To simplify the calculations below we assume $d = 1$, so that $H = H(x, p)$ is only a function of two variables. Let u be a solution of (42). Differentiating (42) in x and applying the chain rule we derive,

$$\partial_t \partial_x u + \partial_p H(x, \partial_x u) \partial_x^2 u = -\partial_x H(x, \partial_x u) \quad (44)$$

Now take $x(t)$ a solution of the equation $\frac{dx}{dt} = \partial_p H(x(t), \partial_x u(x(t)))$ and set $p(t) := \partial_x u(t, x(t))$. Then, by using first the chain rule and then equation (44) we derive,

$$\begin{aligned} \frac{dp}{dt} &= \partial_x \partial_t u(t, x(t)) + \partial_x^2 u(t, x(t)) \partial_p H(x(t), p(t)) \\ &= -\partial_x H(x(t), \partial_x u(t, x(t))) = -\partial_x H(x(t), p(t)) \end{aligned}$$

Hence $x(t), p(t)$ verify the Hamilton equation

$$\frac{dx}{dt} = \partial_p H(x(t), p(t)), \quad \frac{dp}{dt} = -\partial_x H(x(t), p(t)).$$

On the other hand, $\frac{d}{dt}u(t, x(t)) = \partial_t u(t, x(t)) + \partial_x u(t, x(t))\partial_p H(x(t), p(t))$, and, using equation (42), $\partial_t u(t, x(t)) = -H(x(t), p(t))$, $\partial_x u(t, x(t)) = -\partial_p H(x(t), p(t))$. Thus,

$$\frac{d}{dt}u(t, x(t)) = -H(x(t), p(t)) + p(t)\partial_p H(x(t), p(t)),$$

from which we see, in principle, how to construct u based only on the knowledge of the solutions $x(t), p(t)$, called the *bicharacteristic curves* of the nonlinear PDE. Once more singularities can only propagate along bicharacteristics. As in the case of the Burger equation singularities will occur, for essentially, all smooth data; thus a classical, i.e. continuously differentiable, solution can only be constructed locally in time. Both Hamilton-Jacobi equation and hamiltonian systems play a fundamental role in Classical Mechanics as well as in the theory of propagation of singularities in linear PDE. The deep connection between hamiltonian systems and first order Hamilton-Jacobi equations have played an important role in the introduction of the Schrödinger equation in quantum mechanics.

2. *Initial value problem for ODE.* To go further with our general presentation we need to discuss the initial value problem. For simplicity let us start with a first order ODE

$$\partial_x u(x) = f(x, u(x)) \quad (45)$$

subject to the initial condition

$$u(x_0) = u_0 \quad (46)$$

The reader may assume, for simplicity, that (45) is a scalar equation and that f is a nice function of x and u , such as $f(x, u) = u^3 - u + 1 + \sin x$. Observe that the knowledge of the initial data u_0 allows us to determine $\partial_x u(x_0)$. Differentiating the equation (45) with respect to x and applying the chain rule, we derive,

$$\partial_x^2 u(x) = \partial_x f(x, u(x)) + \partial_u f(x, u(x))\partial_x u(x) = \cos x + 3u^2(x)\partial_x u(x) - \partial_x u(x)$$

Hence, $\partial_x^2 u(x_0) = \partial_x f(x_0, u_0) + \partial_u f(x_0, u_0)\partial_x u_0$ and since $\partial_x u(x_0)$ has already been determined we infer that $\partial_x^2 u(x_0)$ can be explicitly calculated from the initial data u_0 . The calculation also involves the function f as well as its first partial derivatives. Taking higher derivatives of the equation (45) we can recursively determine $\partial_x^3 u(x_0)$, as well as all other higher derivatives of u at x_0 . One can then, in principle, determine $u(x)$ with the help of the Taylor series $u(x) = \sum_{k \geq 0} \frac{1}{k!} \partial_x^k u(x_0)(x - x_0)^k = u(x_0) + \partial_x u(x_0)(x - x_0) + \frac{1}{2!} \partial_x^2 u(x_0)(x - x_0)^2 + \dots$. We say in principle because there is no guarantee that the series converge. There is however a very important theorem, called the *Cauchy-Kowalewski theorem*, which asserts that, if the function f is real analytic, which is certainly the case for our $f(x, u) = u^3 - u + 1 + \sin x$, then there exists a neighborhood J of x_0 where the Taylor series converge to a real analytic solution u of the equation. One can easily show that the solution such obtained is the unique solution to (45) subject to the initial condition (46).

The same result may not hold true if we consider a more general equation of the form,

$$a(x, u(x))\partial_x u = f(x, u(x)), \quad u(x_0) = u_0 \quad (47)$$

Indeed the recursive argument outlined above breaks down in the case of the scalar equation $(x - x_0)\partial_x u = f(x, u)$ for the simple reason that we cannot even determine

$\partial_x u(x_0)$ from the initial condition $u(x_0) = u_0$. A similar problem occurs for the equation $(u - u_0)\partial_x u = f(x, u)$. An obvious condition which allows us to extend our previous recursive argument to (47) is that $a(x_0, u_0) \neq 0$. Otherwise we say that the initial value problem (47) is *characteristic*. If both a and f are also real analytic the Cauchy-Kowalewski theorem applies and we obtain a unique, real analytic, solution of (47) in a small neighborhood of x_0 . In the case of a $N \times N$ system,

$$A(x, u(x))\partial_x u = F(x, u(x)), \quad u(x_0) = u_0 \quad (48)$$

$A = A(x, u)$ is $N \times N$ matrix and the *non-characteristic condition* becomes

$$\det A(x_0, u_0) \neq 0. \quad (49)$$

It turns out, and this is extremely important, that while the non-degeneracy condition (49) is essential to obtain a unique solution of the equation, the analyticity condition is not at all important, in the case of ODE. It can be replaced by a simple *local Lipschitz condition* for A and F , i.e. it suffices to assume, for example, that only their first partial derivatives exist and that they are merely locally bounded. This is always the case if the first derivatives of A, F are continuous.

The following local existence and uniqueness (LEU) theorem is called the fundamental theorem of ODE.

Theorem[LEU] *If the matrix $A(x_0, u_0)$ is invertible and if A, F are continuous and have locally bounded first derivatives then there exists a time interval $x_0 \in J \subset \mathbb{R}$ and a unique solution⁷ u defined on J verifying the initial conditions $u(x_0) = u_0$.*

The proof of the theorem is based on the *Picard iteration method*. The idea is to construct a sequence of approximate solutions $u_{(n)}(x)$ which converge to the desired solution. Without loss of generality we can assume A to be the identity matrix⁸. One starts by setting $u_{(0)}(x) = u_0$ and then defines recursively,

$$\partial_x u_{(n)}(x) = F(x, u_{(n-1)}(x)), \quad u_{(n-1)}(x_0) = u_0 \quad (50)$$

Observe that at every stage we only need to solve a very simple linear problem, which makes Picard iteration easy to implement numerically. As we shall see below, variations of this method are also used for solving nonlinear PDE.

Remark. The local existence theorem is sharp, in general. Indeed we have seen that the invertibility condition for $A(x_0, u_0)$ is necessary. Also, in general, the interval of existence J may not be extended to the whole real line. As an example consider the nonlinear equation $\partial_x u = u^2$ with initial data $u = u_0$ at $x = 0$, for which the solution $u = \frac{u_0}{1 - xu_0}$ becomes infinite in finite time, i.e. it *blows-up*.

Once the LEU result is established one can define the main goals of the mathematical theory of ODE to be:

⁷Since we are not assuming analyticity for A, F the solution may not be analytic, but it has continuous first derivatives.

⁸since A is invertible we can multiply both sides of the equation by the inverse matrix A^{-1}

- (1) Find criteria for global existence. In case of blow-up describe the limiting behavior.
- (2) In case of global existence describe the asymptotic behavior of solutions and family of solutions.

Though is impossible to develop a general theory, answering both goals (in practice one is forced to restrict to special classes of equations motivated by applications), the general LEU theorem mentioned above gives a powerful unifying theme. It would be very helpful if a similar situation were to hold for general PDE.

3. Initial value problem for PDE. By analogy to the one dimensional situation it is natural to consider, instead of points, hyper-surfaces $\mathcal{H} \subset \mathbb{R}^d$ on which to specify initial conditions for u . For a general equation of order k , i.e. involving k derivatives, we would need to specify the values of u and its first $k - 1$ normal derivatives⁹ to \mathcal{H} . For example in the case of the second order wave equation (3) we need to specify the initial data for u and $\partial_t u$. along the hypersurface $t = 0$. Without getting into details at this point we can give the following general definition.

Definition. *We say that an initial value problem, for a k -order quasilinear system, in which we specify, as data, the first $k - 1$ normal derivatives of a solution u along \mathcal{H} , is non-characteristic at a point x_0 of \mathcal{H} , if we can formally determine all other higher partial derivatives of u at x_0 , uniquely, in terms of the data.*

To understand the definition, which may seem too general at this point, consider the much simpler case $k = 1$, $N = 1$. In this case we only need to specify the restriction $u|_{\mathcal{H}} = u_0$ of u to \mathcal{H} . Our initial value problem takes the form,

$$\sum_{i=1}^d a^i(x, u(x)) \partial_i u(x) = f(x, u(x)), \quad u|_{\mathcal{H}} = u_0 \quad (51)$$

with a^i , f real valued functions of $x \in \mathbb{R}^d$ and $u \in \mathbb{R}$. To simplify further take $d = 2$, i.e. we have the equation in $x = (x^1, x^2)$,

$$a^1(x, u(x)) \partial_1 u(x) + a^2(x, u(x)) \partial_2 u(x) = f(x, u(x)) \quad (52)$$

we have encountered earlier in (37). Consider a curve \mathcal{H} in \mathbb{R}^2 , parametrized by $x^1 = x^1(s)$, $x^2 = x^2(s)$ whose tangent vector $V(s) = (\frac{dx^1}{ds}, \frac{dx^2}{ds})$ is non-degenerate, i.e. $|V(s)| = (|\frac{dx^1}{ds}|^2 + |\frac{dx^2}{ds}|^2)^{1/2} \neq 0$. It has a well defined unit normal $N(s) = (n_1(s), n_2(s))$, which verifies the conditions,

$$N(s) \cdot V(s) = 0, \quad N(s) \cdot N(s) = 1$$

Observe that the coefficients a^1, a^2 in (52) can be completely determined, along \mathcal{H} , from the knowledge of the initial condition $u_0 = u_0(s)$. Consider the first derivatives $(\partial_1 u, \partial_2 u)$ evaluated along \mathcal{H} , i.e. $U(s) = (\partial_1 u(x(s)), \partial_2 u(x(s)))$. At every point along \mathcal{H} our equation reads,

$$A(s) \cdot U(s) = f(s), \quad (53)$$

⁹These are derivatives in the direction of the normal to \mathcal{H} .

where $A(s) = (a^1(x(s), u_0(s)), a^2(x(s), u_0(s)))$ and $f(s) = f(x(s), u_0(s))$ are completely determined by the data $u_0(s)$. Differentiating $u(x(s)) = u_0(s)$ with respect to s we infer that,

$$U(s) \cdot V(s) = U_0(s), \quad U_0(s) = \frac{d}{ds} u_0(s).$$

To fully determine $U(s)$ it remains to determine its projection on the normal vector $N(s)$, i.e. $U(s) \cdot N(s)$. Indeed, since $V(x)$ and $N(x)$ span \mathbb{R}^2 , at all points $x = (x^1(s), x^2(s))$ along our curve, we have

$$U(s) = (U \cdot V)(s) \frac{V(s)}{|V(s)|^2} + (U \cdot N)(s)N(s) \quad (54)$$

Therefore, from the equation (53),

$$f(s) = A(s) \cdot U(s) = (U(s) \cdot V(s)) \frac{A(s) \cdot V(s)}{|V(s)|^2} + (U(s) \cdot N(s)) A(s) \cdot N(s)$$

from which we can determine $U(s) \cdot N(s)$ provided that,

$$A(s) \cdot N(s) \neq 0. \quad (55)$$

If, on the other hand, $A(s) \cdot N(s) = 0$ then, since $V(s) \cdot N(s) = 0$, we infer that the vectors $A(s)$ and $V(s) = \frac{dx}{ds}$ must be proportional, i.e. $\frac{dx}{ds} = \lambda(s)A(s)$. One can then reparametrize the curve \mathcal{H} , i.e. introduce another parameter $s' = s'(s)$ with $\frac{ds'}{ds} = \lambda(s)$, such that relative to the new parameter we have $\lambda = 1$. This leads to the equation,

$$\frac{dx^1}{ds} = a^1(x(s), u(x(s))), \quad \frac{dx^2}{ds} = a^2(x(s), u(x(s)))$$

which is precisely the characteristic system (38). Thus,

Along a characteristic curve, the equation (52) is degenerate, that is we cannot determine the first order derivatives of u uniquely in terms of the data u_0 . On the other hand the non-degenerate condition,

$$A(s_0) \cdot N(s_0) \neq 0, \quad \text{i.e.} \quad a^1(x_0, u(x_0))n_1(x_0) + a^2(x_0, u(x_0))n_2(x_0) \neq 0 \quad (56)$$

at some point $x_0 = x(s_0) \in \mathcal{H}$, allows us to determine all higher derivatives of u at x_0 , uniquely in terms of the data u_0 .

Indeed, if the condition $A(s_0) \cdot N(s_0) \neq 0$ is satisfied at $x_0 = x(s_0) \in \mathcal{H}$ we have seen already how to determine the first derivatives $\partial_1 u, \partial_2 u$ at that point. Once we have these it is not difficult to determine all higher derivatives of u . For example, observe, by differentiating equation (52) with respect to x^1 , that the function $v = \partial_1 u$ verifies an equation of the form,

$$a^1(x, u(x))\partial_1 v(x) + a^2(x, u(x))\partial_2 v = g(x, u(x), v(x))$$

with a function g which can be easily determined from the coefficients a and f . We can now proceed as before and determine the first derivatives of v i.e. $\partial_1^2 u, \partial_2 \partial_1 u$. Thus, recursively, we can determine all partial derivatives of u of any order.

We can easily extend the discussion above to the higher dimensional case (51). Given a hypersurface \mathcal{H} in \mathbb{R}^d , with unit normal $N = (n_1, n_2, \dots, n_d)$, we find that

\mathcal{H} is non-characteristic at x_0 for the initial value problem (51) if,

$$\sum_{i=1}^n a^i(x_0, u_0(x_0))n_i(x_0) \neq 0 \quad (57)$$

With a little more work we can extend our discussion to general higher order quasi-linear equations, or systems and get a simple, sufficient condition, for a Cauchy problem to be non-characteristic. Particularly important for us are second order ($k = 2$) scalar equations ($N = 1$). To keep things simple consider the case of a second order, semi-linear equation in \mathbb{R}^d ,

$$\sum_{i,j=1}^d a^{ij}(x)\partial_i\partial_j u = f(x, u(x), \partial u(x)) \quad (58)$$

and a hypersurface \mathcal{H} in \mathbb{R}^d defined by the equation $\psi(x) = 0$ with non-vanishing gradient $\partial\psi$. Define the unit normal at a point $x_0 \in \mathcal{H}$ to be $N = \frac{\partial\psi}{|\partial\psi|}$, or in components $n_i = \frac{\partial_i\psi}{|\partial\psi|}$. As initial conditions for (58) we prescribe u and its normal derivative $Nu(x) = n_1(x)\partial_1 u(x) + n_2(x)\partial_2 u(x) + \dots + n_d(x)\partial_d u(x)$ on \mathcal{H} ,

$$u(x) = u_0(x), \quad Nu(x) = u_1(x), \quad x \in \mathcal{H} \quad (59)$$

We need to find a condition on \mathcal{H} such that we can determine all higher derivatives of a solution u , at $x_0 \in \mathcal{H}$, from the initial data u_0, u_1 . We can proceed exactly in the same manner as before, and find that all second order derivatives of u can be determined at a point $x_0 \in \mathcal{H}$, provided that,

$$\sum_{i,j=1}^d a^{ij}(x_0)n_i(x_0)n_j(x_0) \neq 0 \quad (60)$$

It is indeed easy to see that the only second order derivative of u , which is not automatically determined from u_0, u_1 , is of the form $N^2 u(x_0) = N(N(u))(x_0)$. This latter can be determined from the equation (58), provided that (60) is verified. One does this by decomposing all partial derivatives of u into tangential and normal components, as we have done in (54). One can then show, recursively, that all higher derivatives of u can also be determined. Thus, (60) is exactly the *non-characteristic* condition we were looking for.

If, on the other hand, $\sum_{i,j=1}^d a^{ij}(x)n_i(x)n_j(x) = 0$ at all points we call \mathcal{H} a *characteristic hypersurface* for the equation (58). Since $n_i = \frac{\partial_i\psi}{|\partial\psi|}$ we find that \mathcal{H} is characteristic if and only if,

$$\sum_{i,j=1}^d a^{ij}(x)\partial_i\psi(x)\partial_j\psi(x) = 0 \quad (61)$$

Example 1. Assume that the coefficients a of (58) verify,

$$\sum_{i,j=1}^d a^{ij}(x)\xi_i\xi_j > 0, \quad \forall \xi \in \mathbb{R}^d, \quad \forall x \in \mathbb{R}^d \quad (62)$$

Then no surface in \mathbb{R}^d can be characteristic. This is the case, in particular, for the equation $\Delta u = f$. Consider also the minimal surfaces equation written in the form (32). It is easy to check that, the quadratic form associated to the symmetric matrix $h^{ij}(\partial u)$ is positive definite independent of ∂u . Indeed,

$$h^{ij}(\partial u)\xi_i\xi_j = (1 + |\partial u|^2)^{-1/2}(|\xi|^2 - (1 + |\partial u|^2)^{-1}(\xi \cdot \partial u)^2) > 0$$

Thus, even though (32) is not linear, we see that all surfaces in \mathbb{R}^2 are non-characteristic.

Example 2. Consider the wave equation $\square u = f$ in \mathbb{R}^{1+d} . All hypersurfaces of the form $\psi(t, x) = 0$ for which,

$$(\partial_t \psi)^2 = \sum_{i=1}^d (\partial_i \psi)^2, \tag{63}$$

are characteristic. This is the famous Eikonal equation which plays a fundamental role in the study of wave propagation. Observe that it splits into two Hamilton-Jacobi equations, see (42),

$$\partial_t \psi = \pm \left(\sum_{i=1}^d (\partial_i \psi)^2 \right)^{1/2} \tag{64}$$

The bicharacteristic curves of the associated Hamiltonians are called bicharacteristic curves of the wave equation. As particular solutions of (63) we find, $\psi_+(t, x) = (t - t_0) + |x - x_0|$ and $\psi_-(t, x) = (t - t_0) - |x - x_0|$ whose level surfaces $\psi_{\pm} = 0$ correspond to forward and backward light cones with vertex at $p = (t_0, x_0)$. These represent, physically, the union of *all light rays emanating from a point source at p*. The light rays are given by the equation $(t - t_0)\omega = (x - x_0)$, for $\omega \in \mathbb{R}^3$ with $|\omega| = 1$, and are precisely the (t, x) components of the bicharacteristic curves of the Hamilton-Jacobi equations (64). More general, the characteristics of the linear wave equation,

$$a^{00}(t, x)\partial_t^2 u - \sum_{i,j} a^{ij}(t, x)\partial_i\partial_j u = 0, \tag{65}$$

with $a^{00} > 0$ and a^{ij} verifying (62), are given by the Hamilton-Jacobi equations,

$$-a^{00}(t, x)(\partial_t \psi)^2 + a^{ij}(x)\partial_i \psi \partial_j \psi = 0 \tag{66}$$

or,

$$\partial_t \psi = \pm \left((a^{00})^{-1} \sum_{i,j} a^{ij}(x)\partial_i \psi \partial_j \psi \right)^{1/2}. \tag{67}$$

The bicharacteristics of the corresponding hamiltonian systems are called *bicharacteristic curves* of (65).

Remark. In the case of the first order scalar equations (35) we have seen how the knowledge of characteristics can be used to find, implicitly, the general solutions. We have shown, in particular, that singularities propagate only along characteristics. In the case of second order equations the characteristics are not sufficient to solve the equations, but they continue to provide important information, such as propagation of singularities. For example, in the case of the wave equation $\square u = 0$

with smooth initial data u_0, u_1 everywhere except at a point $p = (t_0, x_0)$, the solution u has singularities present at all points of the light cone $-(t-t_0)^2 + |x-x_0|^2 = 0$ with vertex at p . A more refined version of this fact shows that the singularities propagate along bicharacteristics. The general principle here is that *singularities propagate along characteristic hypersurfaces of a PDE*. Since this is a very important principle it pays to give it a more precise formulation which extends to general boundary conditions, such as the Dirichlet condition for (1).

Propagation of singularities¹⁰. *If the boundary conditions, or the coefficients of a PDE are singular at some point p , and smooth (or real analytic) away from p in some small neighborhood V , then a solution of the equation may only be singular in V along a characteristic hypersurface passing through p . If there are no such characteristic hypersurfaces, any solution of the equation must be smooth (or real analytic) in $V \setminus \{p\}$.*

Remark 1. The heuristic principle mentioned above is invalid, in general, in the large. Indeed, as we have shown in the case of the Burger equation, solutions to nonlinear evolution equations, can develop new singularities independent of the smoothness of the initial conditions. Global versions of the principle can be formulated for linear equations, based on the bicharacteristics of the equation, see remark 3 below.

Remark 2. According to the principle it follows that any solution of the equation $\Delta u = f$, verifying the boundary condition $u|_{\partial D} = u_0$, with a boundary value u_0 which is merely continuous, has to be smooth everywhere in the interior of D provided that f itself is smooth there. Moreover the solution is real analytic, if f is real analytic.

Remark 3. More precise versions of this principle, which plays a fundamental role in the general theory, can be given for linear equations. In the case of the general wave equation (65), for example, one can show that singularities propagate along bicharacteristics. These are the bicharacteristic curves associated to the Hamilton-Jacobi equation (67).

3. *Cauchy-Kowalevsky.* In the case of ODE we have seen that a non-characteristic initial value problem admits always local in time solutions. Is there also a higher dimensional analogue of this fact? The answer is yes provided that we restrict ourselves to an extension of the Cauchy-Kowalevsky theorem. More precisely one can consider general quasilinear equations, or systems, with real analytic coefficients, real analytic hyper-surfaces \mathcal{H} , and real analytic initial data on \mathcal{H} .

Theorem[Cauchy-Kowalevsky (CK)] *If all the real analyticity conditions made above are satisfied and if \mathcal{H} is non-characteristic at x_0 ¹¹, there exists locally, in a neighborhood of x_0 , a unique real analytic solution $u(x)$ verifying the system and the corresponding initial conditions.*

¹⁰A more precise version of the principle relates propagation of singularities to bicharacteristics curves.

¹¹In the case of second order equations of type (58) this is precisely condition (60).

In the special case of linear equations (33) an important companion theorem, due to Holmgren, asserts that the analytic solution given by the CK theorem is unique in the class of all smooth solutions and smooth non-characteristic hypersurfaces \mathcal{H} . The CK theorem validates the most straightforward attempts to find solutions by formal expansions $u(x) = \sum_{\alpha} C_{\alpha}(x - x_0)^{\alpha}$ with constants C_{α} which can be determined recursively, by simply algebraic formulas, from the equation and initial conditions on \mathcal{H} , using only the non-characteristic condition and the analyticity assumptions. Indeed the theorem insures that the naive expansion obtained in this way converges in a small neighborhood of $x_0 \in \mathcal{H}$.

It turns out, however, that the analyticity conditions required by the CK theorem are much too restrictive and thus the apparent generality of the result is misleading. A first limitation becomes immediately obvious when we consider the wave equation $\square u = 0$ whose fundamental feature of finite speed of propagation¹² (see companion article) is impossible to make sense in the class of real analytic solutions. A related problem, first pointed out by Hadamard, concerns the impossibility of solving the Cauchy problem, in many important cases, for arbitrary smooth, non analytic, data. Consider, for example, the Laplace equation $\Delta u = 0$ in \mathbb{R}^d . As we have established above, any hyper-surface \mathcal{H} is non-characteristic, yet the Cauchy problem $u|_{\mathcal{H}} = u_0$, $N(u)|_{\mathcal{H}} = u_1$, for arbitrary smooth initial conditions u_0, u_1 may admit no local solutions, in a neighborhood of any point of \mathcal{H} . Indeed take \mathcal{H} to be the hyperplane $x_1 = 0$ and assume that the Cauchy problem can be solved, for a given, non analytic, smooth data in an domain which includes a closed ball B centered at the origin. The corresponding solution can also be interpreted as the solution to the Dirichlet problem in B , with the values of u prescribed on the boundary ∂B . But this, according to our heuristic principle¹³, must be real analytic everywhere in the interior of B , contradicting our initial data assumptions.

On the other hand the Cauchy problem, for the wave equation $\square u = 0$ in \mathbb{R}^{d+1} , has a unique solution for any smooth initial data u_0, u_1 , prescribed on a *space-like hyper-surface*, that is a hypersurface $\psi(t, x) = 0$ whose normal vector, at every point $p = (t_0, x_0)$, is directed inside the interior of the future or past directed light cone passing through that point. Analytically this means,

$$|\partial_t \psi(p)| > \left(\sum_{i=1}^d |\partial_i \psi(p)|^2 \right)^{1/2}. \quad (68)$$

The condition is clearly satisfied by the hypersurfaces of $t = t_0$, but any other hypersurface close to it is also spacelike. On the other hand the IVP is *ill posed*, i.e. not well posed, for a time-like hypersurface, i.e a hypersurface for which,

$$|\partial_t \psi(p)| < \left(\sum_{i=1}^d |\partial_i \psi(p)|^2 \right)^{1/2}. \quad (69)$$

¹²Roughly this means that if a solution u is compactly supported at some value of t it must be compactly supported at all later times. Analytic functions cannot be compactly supported without vanishing identically.

¹³which can be easily made rigorous in this case

In this case we cannot, for general non real analytic initial conditions, find a solution of the IVP. An example of a time-like hypersurface is given by the hyperplane $x^1 = 0$.

Definition. *A given problem for a PDE is said to be well posed if both existence and uniqueness of solutions can be established for arbitrary data which belong to a specified large space of functions, which includes the class of smooth functions¹⁴. Moreover the solutions must depend continuously on the data.*

The continuous dependence on the data is very important. Indeed the IVP would be of little use if very small changes of the initial conditions will result, instantaneously, in very large changes in the corresponding solutions.

4. *Standard classification.* The different behavior of the Laplace and Wave equations mentioned above illustrates the fundamental difference between ODE and PDE and the illusory generality of the CK theorem. Given that the Laplace and wave equation are so important in geometric and physical application one is interested to find the broadest classes of equations with which they share their main properties. The equations modeled by the Laplace equation are called *elliptic* while those modeled by the wave equation are called *hyperbolic*. The other two important models are the the heat, see (2), and Schrödinger equation, see (5). The general classes of equations with which they resemble are called *parabolic* and, respectively, *dispersive*.

Elliptic equations are the most robust and easiest to characterize, they admit no characteristic hypersurfaces.

Definition 1: *A linear, or quasi-linear, $N \times N$ system with no characteristic hyper-surfaces is called elliptic.*

Clearly the equations of type (58) whose coefficients a^{ij} verify condition (62) are elliptic. The minimal surface equation (6) is also elliptic. It is also easy to verify that the Cauchy-Riemann system (13) is elliptic. As it was pointed out by Hadamard, the initial value problem is not well posed for elliptic equations. The natural way of parametrizing the set of solutions to an elliptic PDE is to prescribe conditions for u , and some of its derivatives¹⁵, at the boundary of a domain $D \subset \mathbb{R}^n$. These are called *boundary value problems (BVP)*. A typical example is the Dirichlet boundary condition $u|_{\partial D} = u_0$ for the Laplace equations $\Delta u = 0$ in a domain $D \subset \mathbb{R}^n$. One can show that, under mild regularity assumptions on the domain D and continuous boundary value u_0 , this problem admits a unique solution, depending continuously on u_0 . We say that the Dirichlet problem for the Laplace equation is well posed. Another well posed problem for the Laplace equation is given by the Neumann boundary condition $N(u)|_{\partial D} = f$, with N the exterior unit normal to the boundary. The problem is well posed for all continuous functions f defined on ∂D with zero mean average. A typical problem of general theory is to classify all well posed BVP for a given elliptic system.

¹⁴Here we are necessarily vague. A precise space can be specified in each specific case.

¹⁵roughly half the order of the equation

As a consequence of our propagation of singularities principle, we deduce, heuristically, the following general fact:

Classical solutions of elliptic equations with smooth (or real analytic) coefficients in a regular domain D are smooth (or real analytic), in the interior of D , independent of how smooth the boundary conditions¹⁶ are.

Hyperbolic equations are, essentially, those for which the initial value problem is well posed. In that sense, they provide the natural framework for which one can prove a result similar to the local existence theorem for ODE. More precisely, for each sufficiently regular set of initial conditions there corresponds a unique solution; we can thus think of the Cauchy problem as a natural way of parametrizing the set of all solutions to the equations.

The definition of hyperbolicity depends, however, on the particular hypersurface we are considering. Thus, in the case of the wave equation $\square u = 0$, the standard initial value problem

$$u(0, x) = u_0(x), \quad \partial_t u(0, x) = u_1$$

is well posed. This means that for any smooth initial data u_0, u_1 we can find a unique solution of the equation which depends continuously on u_0, u_1 . As we have mentioned earlier, the IVP for $\square u = 0$ remains well posed if we replace the initial hypersurface $t = 0$ by any space-like hypersurface $\psi(t, x) = 0$, see (68). It fails however to be well posed for timelike hypersurfaces, see (69). In that case there may not exist any solution with prescribed, non-analytic, Cauchy data.

It is more difficult to give algebraic conditions of hyperbolicity. In principle hyperbolic equations differ from the elliptic ones, roughly, by the presence of a maximum number of characteristic hypersurfaces passing through any given point. Rather than attempting a general definition is more useful to give some examples.

One of the most useful class of hyperbolic equations is given by second order wave equations of the form

$$\square_a u = f, \quad \square_a = -a^{00} \partial_t^2 + \sum_{i,j=1}^d a^{ij} \partial_i \partial_j \quad (70)$$

with coefficients a^{00} , a^{ij} and f which may depend on (t, x) as well as $u(t, x)$ and $\partial u(t, x)$. We need also to assume that $a^{00} > 0$ and a^{ij} verify the ellipticity condition,

$$\sum_{i,j=1}^d a^{ij} \xi_i \xi_j > 0, \quad \xi \in \mathbb{R}^d \quad (71)$$

¹⁶Provided that the boundary condition under consideration is well posed. Moreover this heuristic principle holds, in general, only for classical solutions of a nonlinear equation. There are in fact examples of well posed boundary value problems, for nonlinear elliptic systems, with no classical solutions.

The IVP for this type of equations is well posed, for any hypersurface $\psi(t, x) = 0$, such as $t = t_0$, for which,

$$-a^{00}(\partial_t \psi)^2 + \sum_{i,j=1}^d a^{ij} \partial_i \psi \partial_j \psi < 0, \quad \xi \in \mathbb{R}^d \quad (72)$$

A very useful generalization of (70) consist of the class of system of wave equation, diagonal with respect to the second derivatives, i.e. ,

$$\square_a u^I = f^I(u, \partial u), \quad I = 1, 2, \dots, N, \quad (73)$$

where $u = (u^1, u^2, \dots, u^N)$. One can check, see (22), that the Einstein equations, in wave coordinates, can be written, almost, in this form.

Remark In reality (22) are slightly different¹⁷. The operator \square_a has to be replaced by an operator of the form $\square_g = g^{\alpha\beta} \partial_\alpha \partial_\beta$ where $g^{\alpha\beta}$ is the inverse of a general Lorentz metric $g_{\alpha\beta}$. The characteristics surfaces of the wave operator \square_g , called also *null hypersurfaces of the metric g*, are given by the equation, to be compared with (66),

$$g^{\alpha\beta} \partial_\alpha \psi \partial_\beta \psi = 0. \quad (74)$$

Another important class, which includes most of the important known examples of first order hyperbolic equations, such as Maxwell, are of the form,

$$A^0(t, x, u) \partial_t u + \sum_{i=1}^d A_i(t, x, u) \partial_i u = F(t, x, u), \quad u|_{\mathcal{H}} = u_0 \quad (75)$$

where all the coefficients A^0, A^1, \dots, A^d are symmetric $N \times N$ matrices and \mathcal{H} is given by $\psi(t, x) = 0$. Such a system is well posed provided that the matrix,

$$A^0(t, x, u) \partial_t \psi(t, x) + \sum_{i=1}^d A_i(t, x, u) \partial_i \psi(t, x) \quad (76)$$

is positive definite. A system (75) verifying these conditions is called *symmetric hyperbolic*. In the particular case when $\psi = t$ the condition (76) becomes

$$(A^0 \xi, \xi) \geq c |\xi|^2 \quad \forall \xi \in \mathbb{R}^N. \quad (77)$$

Remark. It turns out that the second order wave equation¹⁸ (70), verifying (72) can be written as a first order symmetric hyperbolic system. This can be simply done by introducing the new variables $v_0 = \partial_t u$, $v_1 = \partial_1 u, \dots, v_d = \partial_d u$ and the obvious compatibility relations, such as $\partial_i v_j = \partial_j v_i$.

The following is a fundamental result in the theory of general hyperbolic equations It is called the local existence and uniqueness [LEU] for symmetric hyperbolic systems:

Theorem[LEU-Hyperbolic] *The initial value problem (75), is locally well posed, for symmetric hyperbolic systems, with sufficiently smooth A, F, \mathcal{H} and sufficiently*

¹⁷It turns out however that, locally, one can always change coordinates such that the components $g^{0i}, i = 1, 2, \dots, d$ vanish identically

¹⁸as well as the diagonal system (73)

smooth initial conditions u_0 . In other words, if the above conditions are satisfied, then for any point $p \in \mathcal{H}$ there exist a sufficiently small neighborhood $\mathcal{D} \subset \mathbb{R}^{1+d}$ of p and a unique, continuously differentiable, solution $u : \mathcal{D} \rightarrow \mathbb{R}^N$.

Remark 1. The issue of how smooth the initial data is allowed to be is an important question, still under investigation, for nonlinear equations.

Remark 2. The local character of the theorem is essential, the result cannot be globally true, in general for nonlinear systems. Indeed, as we have seen, the evolution problem (39) for the Burger equation, which fits trivially into the framework of symmetric hyperbolic systems, leads, after a sufficiently large time, to singular solutions. This happens independent of how smooth the initial data u_0 is. A precise version of the theorem above gives a lower bound on how large \mathcal{D} can be.

Remark 3. The proof of the theorem is based on a variation of the Picard iteration method we have encountered earlier for ODE. One starts by taking $u_{(0)} = u_0$ in a neighborhood of \mathcal{H} and then define recursively,

$$A^0(t, x, u_{(n-1)})\partial_t u_{(n)} + \sum_{i=1}^d A_i(t, x, u_{(n-1)})\partial_i u_{(n)} = F(t, x, u_{(n-1)}), \quad u_{(n)}|_{\mathcal{H}} = u_0 \quad (78)$$

Observe that at every stage of the iteration we have to solve a linear equation. Linearization is an extremely important tool in studying nonlinear PDE. We can almost never understand their behavior without linearizing them around important special solutions. Thus, almost invariably, hard problems in non-linear PDE reduce to understanding specific problems in linear PDE.

Remark 4. To implement the Picard iteration method we need to get precise estimates on the u_n iterate in terms of the u_{n-1} iterate. This step requires energy type a-priori estimates which we discuss in the next section.

Remark 5. Theorem[LEU-hyperbolic] has wide applications to various hyperbolic systems of physical interests. It applies, in particular, to prove a local existence result for the Einstein equations in wave coordinates, see (22).

Another important, characteristic¹⁹, property of hyperbolic equations is *finite speed of propagation*. Consider the simple case of the wave equation (3). In this case the initial value problem can be solved explicitly by the Kirchoff formula (see companion article). The formula allows us to conclude that if the initial data, at $t = 0$, is supported in a ball $B_a(x_0)$ of radius $a > 0$ centered at $x_0 \in \mathbb{R}^3$ then at time $t > 0$ the solution u is supported in the ball $B_{a+t}(x_0)$. In general finite speed of propagation can be best formulated in terms of domains of dependence and influence of hyperbolic equations. Given a point $p \in \mathbb{R}^{1+d}$, outside the initial hypersurface \mathcal{H} , we define $\mathcal{D}(p) \subset \mathcal{H}$ as the complement of the set of points $q \in \mathcal{H}$ with the property that any change of the initial conditions made in a small neighborhood V of q does not influence the value of solutions at p . More precisely if u, v are two solutions of the equation whose initial data differ only in V , must also coincide at p . The property of finite speed of propagation simply means that, for any point

¹⁹Elliptic, parabolic and dispersive equations do not have this property.

p , $\mathcal{D}(p)$ is compact in \mathcal{H} . A related notion is that of *domain of influence*. Given a set $D \subset \mathcal{H}$ the domain of influence of D is the smallest set $\mathcal{J}(D) \subset \mathbb{R}^{1+d}$ with the property that any two solutions u, v of the equation whose initial conditions coincide in the complement of D , must also coincide at all points in the complement of $\mathcal{J}(D)$. In the case of $\square u = 0$, if at $t = 0$, u and $\partial_t u$ are zero outside the unit ball B , $|x| \leq 1$, then, u is identically zero in the region $|x| > 1 + |t|$. Thus $\mathcal{J}(B)$ must be a subset of $\{(t, x) / |x| \leq 1 + |t|\}$ and it can be shown that in fact $\mathcal{J}(B) = \{(t, x) / |x| \leq 1 + |t|\}$. Observe also that the boundary of $\mathcal{J}(B)$ is formed by the union of two smooth characteristic hypersurfaces of the wave equation, $|x| = t + 1$ for $t \geq 0$ and $|x| = -t + 1$ for $t \leq 0$. This is a general fact, which illustrates once more the importance of characteristics.

The boundaries of domains of dependence of classical solutions to hyperbolic PDE are characteristic hypersurfaces, typically piecewise smooth.

Finally a few words for parabolic equations and Schrödinger type equations²⁰. A large class of useful equations of this type is given by,

$$\partial_t u - Lu = f, \quad (79)$$

and, respectively

$$i\partial_t u + Lu = f \quad (80)$$

where L is the elliptic operator $L = \sum_{i,j=1}^d a^{ij} \partial_i \partial_j$ verifying the ellipticity condition (71). One looks for solutions $u = u(t, x)$, defined for $t \geq t_0$, with the prescribed initial condition,

$$u(t_0, x) = u_0(x) \quad (81)$$

on the hypersurface $t = t_0$. Strictly speaking this hypersurface is characteristic, since the order of the equation is $k = 2$ and we cannot determine $\partial_t^2 u$ at $t = t_0$ directly from the equation. Yet this is not a serious problem; we can still determine $\partial_t^2 u$ formally by differentiating the equation with respect to ∂_t . Thus, the initial value problem (79), (resp. (80)) and (81) is well posed, but in a slightly different sense than for hyperbolic equations. For example the heat equation $-\partial_t u + \Delta u$ is only well posed for positive t and ill posed for negative t . The heat equation may also not have unique solutions for the IVP unless we make assumptions about how fast the initial data is allowed to grow at infinity. One can also show that the only characteristics of the equation (79) are all of the form $t = t_0$ and therefore parabolic equations are quite similar to elliptic equations. For, example, one can show, consistent with our propagation of singularities principle, that if the coefficients a^{ij} and f are smooth (or real analytic), then, even if the initial data u_0 may not be smooth, the solution u must be smooth (or real analytic in x) for $t > t_0$. The heat equation smoothes out initial conditions. It is for this reason that the heat equation is useful in many applications. One often encounters diagonal systems of parabolic equations, of the form

$$\partial_t u^I - Lu^I = f^I(u, \partial u), \quad u = (u^1, u^2, \dots, u^N)$$

with L as above. The system of equations (20), connected with the Ricci flow, is of this form.

²⁰General classes of dispersive equations are a bit harder to describe.

5. *Linear Equations.* General theory has been most successful in regard to linear equations (33). This is particularly true for linear equations with constant coefficients, for which Fourier analysis provides an extremely powerful tool, and for general, linear, elliptic equations. We also have a reasonably good theory for variable coefficients hyperbolic equations²¹, though less complete as in the elliptic case. The following remarks are useful to understand how elliptic and hyperbolic equations fit relative to other classes of equations.

Remark 1: Elliptic and hyperbolic equations are the most robust, useful, classes of PDE. Other important classes, such as *parabolic* and *dispersive*, can be interpreted as lying at the boundaries²² of these two classes. A neat classification of all linear equations into, elliptic, hyperbolic, parabolic and dispersive is unfortunately not possible, even for second order equations in two space dimensions.

Remark 2: A typical property of linear elliptic equations is local regularity. This means that all solutions of a linear elliptic equation, with smooth coefficients²³, must be smooth in any open domain where the equation is satisfied. This follows, heuristically, from the propagation of singularities principle, discussed earlier, and absence of characteristic surfaces. Parabolic equations share this property. Linear hyperbolic equations, on the other hand, have a radically different behavior. Singularities of the initial data propagate along characteristic hypersurfaces, according to simple propagation laws. Finally, dispersive equations have more complicated propagation of singularities properties.

Remark 3. In the case of linear hyperbolic equations, the characteristics of the equations, which are solutions to nonlinear Hamilton-Jacobi equations, can be used to construct approximate solutions to the equations, called *parametrics*, from which one can read the relevant information concerning propagation of singularities²⁴. For constant coefficient equations these parametrics can in fact be constructed to obtain exact solutions.

Remark 4. As we have pointed out in the previous section, hard problems in non-linear PDE are almost always connected with specific linear problems. Yet, often, the linear problems which arise in this way are rather special and cannot be treated with the degree of generality (and it is not at all necessary that they should be !) one expects from a general theory. We will discuss examples of this type in the next section.

In connection with well-posedness we note that there exist scalar, linear, operators P and smooth functions f for which the equation $Pu = f$ may have *no solutions*, in any domain $\Omega \subset \mathbb{R}^n$. The problem of characterizing linear equations which have

²¹Symmetric hyperbolic systems are suitable well-posedness and finite speed of propagation, but not so useful for the more refined question of propagation of singularities. For this goal one uses instead strictly hyperbolic systems or various definitions of hyperbolic systems of higher multiplicity.

²²parabolic equations are singular, formal, limits of elliptic equations. Dispersive equations can be regarded also as singular limits of hyperbolic equations.

²³Including the inhomogeneous terms of the equations, such as $\Delta u = f$

²⁴which propagate along the bicharacteristics of the associated Hamiltonian

the property of *local solvability* has been extensively studied. Today it remains an important, even though less active, area of research in PDE.

Questions of *unique continuation* of solutions are also investigated by the general theory. Unique continuation results concern *ill posed problems* where general existence may fail, yet uniqueness survives. A typical example is Holmgren's theorem mentioned above. It asserts, in the particular case of the wave equation, that, even though the Cauchy problem for time-like hyper-surfaces is ill posed, if a solution exists it must necessarily be unique. More precisely, assume that a solution u of (3) is such that u and $\partial_z u$ vanish along the hyperplane $z = 0$. Then u must vanish identically in the whole space. Ill posed problems appear naturally in connection to control theory which deals with unphysical, auxiliary, boundary conditions which are introduced to guide solutions of the system to a desired state.

Besides the traditional questions of classification, local and global well-posedness, propagation of singularities, and unique continuation of solutions, there are other issues which are addressed by the general theory of linear PDE. A very active area of investigation is *spectral theory*. There is no way I can even begin to give an account of this theory, which is of fundamental importance not only to Quantum Mechanics, and other physical theories, but also to geometry and analytic number theory. A typical problem in spectral theory is to solve the eigenvalue problem in \mathbb{R}^d , or a domain $D \subset \mathbb{R}^d$,

$$-\Delta u(x) + V(x)u(x) = \lambda u(x) \quad (82)$$

that is to find the values $\lambda \in \mathbb{R}$, called eigenvalues, for which there exist solutions $u(x)$, localized in space, i.e. bounded in the $L^2(\mathbb{R}^d)$ norm, called eigenfunctions. The existence of an eigenfunction u implies that we can write solutions to the Schrödinger equation,

$$i\partial_t \phi + \Delta \phi - V\phi = 0 \quad (83)$$

of the form $\phi(t, x) = e^{-i\lambda t}u(x)$, called bound states of the physical system described by (83). The eigenvalues λ correspond to the quanta energy levels of the system. They are very sensitive to the choice of potential V . The distribution of the eigenvalues of the Laplace operator Δ in a domain $D \subset \mathbb{R}^d$ depends on the geometry of the domain D , this is the case, for example, of the very important Weyl asymptotic formula. The *inverse spectral problem* is also important, can one determine the potential V from the knowledge of the corresponding eigenvalues? The eigenvalue problem can be studied in considerable generality by replacing the operator $-\Delta + V$ with a general elliptic operator. More to the point is the study the eigenvalue problem for the Laplace-Beltrami operator associated to a Riemannian manifold. In the particular case of two dimensional manifolds of constant negative Gauss curvature, i.e. $K = -1$, this problem is important in number theory. A famous problem in differential geometry is to characterize the metric on a 2-dimensional compact manifold, from the spectral properties of the associated Laplace-Beltrami operator.

Related to spectral theory, in a sense opposite to it, is scattering theory. In the case of equation (83) solutions which scatter are those which behave freely as $t \rightarrow \infty$, i.e. they behave like solutions to the free Schrödinger equation $i\partial_t \psi + \Delta \psi = 0$. A typical

problem in scattering theory is to show that, if $V(x)$ tends to zero sufficiently fast as $|x| \rightarrow \infty$, all solutions, except the bound states, scatter as $t \rightarrow \infty$.

5. *Conclusions.* In the analytic case, the CK theorem allows us to solve, locally, the IVP for very general classes of PDE. We have a general theory of characteristic hypersurfaces of PDE and understand in considerable generality how they relate to propagation of singularities. We can also distinguish, in considerable generality, the fundamental classes of elliptic and hyperbolic equations and can define general parabolic and dispersive equations. The IVP for a large class of nonlinear hyperbolic systems can be solved locally in time, for sufficiently smooth initial conditions. Similar, local in time, results hold for general classes of nonlinear parabolic and dispersive equations. A lot more can be done for linear equations. We have satisfactory results concerning regularity of solutions for elliptic and parabolic equations and a good understanding of propagation of singularities for a large class of hyperbolic equations. Some aspects of spectral theory and scattering theory and problems of unique continuation can also be studied in considerable generality.

The main defect of the general theory concerns the passage from local to global. Important global features of special equations are too subtle to fit into a too general scheme; on the contrary each important PDE requires special treatment. This is particularly true for nonlinear equations; the large time behavior of solutions is very sensitive to the special features of the equation at hand. Moreover, general points of view may obscure, through unnecessary technical complications, the main properties of the important special cases. A useful general framework is one which provides a simple and elegant treatment of a particular phenomenon, as is the case of symmetric hyperbolic systems in connection to local well posedness and finite speed of propagation. Yet symmetric hyperbolic systems turn out to be simply too general for the study of more refined questions concerning the important examples of hyperbolic equations.

3. GENERAL IDEAS

As we turn away from the general theory, one may be inclined to accept the pragmatic point of view according to which PDE is not a real subject but rather a collection of subjects, such as Hydrodynamics, General Relativity, Several Complex Variables, Elasticity, etc, each organized around a special equation. This rather widespread view point, which makes it difficult to see PDE as a subject in its own right, has its own serious drawbacks. Even though specific equations have specific properties the tools which are used to derive them are intimately related. There exists, in fact, an impressive body of knowledge relevant to all, or large classes of, important equations. Below is a short list of some of our most important general ideas ²⁵.

²⁵I fail to mention, in the few examples given above some of the important functional analytic tools connected to Hilbert space methods, compactness, the implicit function theorems etc. I also fail to mention the importance of probabilistic methods or the development of topological methods for dealing with global properties of elliptic PDE's.

1) *Well posedness.* As we have stated in the previous section well posed problems are at the heart of the modern theory of PDE. Problems are said to be well-posed if they admit unique solutions for given smooth initial or boundary conditions. The corresponding solutions have to depend continuously on the data. This leads to the classification of PDE into elliptic, hyperbolic, parabolic and dispersive. The first step in the study of a nonlinear evolution equation is a proof of a local in time existence and uniqueness theorem, similar to the one for ODE. *Ill posedness*, the counterpart of well-posedness, is also important in many applications. The Cauchy problem for the wave equation (3), with data on the time-like hypersurface $z = 0$, is a typical example. Ill posed problems appear naturally in Control Theory and Inverse Scattering.

2) *Explicit representations. Fundamental solutions.* The basic equations (2)–(5) can be solved explicitly. For example, the solution to the initial value problem for the heat equation in \mathbb{R}_+^{1+d} , i.e. $t \geq 0$,

$$-\partial_t u + \Delta u = 0, \quad u(0, x) = u_0(x)$$

is given by,

$$u(t, x) = \int_{\mathbb{R}^d} E_d(t, x - y) u_0(y) dy$$

where E_d , defined explicitly by $E_d(t, x) = (4\pi t)^{-d/2} e^{-|x|^2/4t}$ for $t > 0, x \in \mathbb{R}^d$ and $E_d = 0$ for $t \leq 0, x \in \mathbb{R}^d$, is called the *fundamental solution* of the heat operator $-\partial_t + \Delta$. Observe that E_d verifies, the equation $(-\partial_t + \Delta)E = 0$ both in both regions $t < 0$ and $t > 0$ but it has a singularity at $t = 0$ which prevents E to verify the equation in the whole space \mathbb{R}^{1+d} . In fact we can check that for any function²⁶ $\phi \in C_0^\infty(\mathbb{R}^{d+1})$, we have,

$$\int_{\mathbb{R}^{d+1}} E_d(t, x) (\partial_t \phi(t, x) + \Delta \phi(t, x)) dt dx = \phi(0, 0) \quad (84)$$

In the language of distribution theory (see compendium article) formula (84) means that E_d , as a distribution, verifies the equation $(-\partial_t + \Delta)E_d = \delta_0$, with δ_0 the Dirac distribution in \mathbb{R}^{1+d} supported at the origin, i.e., $\delta_0(\phi) = \phi(0, 0)$, $\forall \phi \in C_0^\infty(\mathbb{R}^{d+1})$.

Similarly we can represent solutions to the equation $\Delta \phi = f$ in \mathbb{R}^d , by the formula,

$$\phi = \int_{\mathbb{R}^d} K_d(x - y) f(y) dy$$

where, for $d = 2$, $K_2(x) = (2\pi)^{-1} \log |x|$ while for $d \geq 3$, $K_d(x) = ((2-d)\omega_d)^{-1} |x|^{2-d}$, with ω_d the area of the unit sphere in \mathbb{R}^d . Once more we can check that,

$$\int_{\mathbb{R}^d} K_d(x) \Delta \phi(x) dx = \phi(0), \quad \forall \phi \in C_0^\infty(\mathbb{R}^d) \quad (85)$$

Thus, in the sense of distributions, $\Delta K_d = \delta_0$. A similar notion of fundamental solution can be given both for wave, Klein -Gordon and Schrödinger equations. In all cases, the corresponding fundamental solution is a distribution E which verifies, for the corresponding operator $LE = \delta_0$, for the corresponding operator L .

²⁶that is function which is smooth and has compact support in \mathbb{R}^{1+d}

A powerful method of solving linear PDE with constant coefficients is based on the Fourier transform. For example, consider the heat equation $\partial_t - \Delta u = 0$ in one space dimension, with initial condition $u(0, x) = u_0$. Define $\hat{u}(t, \xi)$ to be the Fourier transform of u relative to the space variable,

$$\hat{u}(t, \xi) = \int_{-\infty}^{+\infty} e^{-ix\xi} u(t, x) dx.$$

It is easy to see that $\hat{u}(t, \xi)$ verifies the differential equation,

$$\partial_t \hat{u}(t, \xi) = -\xi^2 \hat{u}(t, \xi), \quad \hat{u}(0, \xi) = \hat{u}_0(\xi)$$

which can be solved by a simple integration and derive, $\hat{u}(t, \xi) = \hat{u}_0(\xi) e^{-t|\xi|^2}$. Thus, using the inverse Fourier transform, we derive a formula for $u(t, x)$,

$$u(t, x) = (2\pi)^{-1} \int_{-\infty}^{+\infty} e^{ix\xi} e^{-t|\xi|^2} \hat{u}_0(\xi) d\xi \tag{86}$$

Similar formulas can be derived for our other basic evolution equations. For example, in the case of the wave equation $-\partial_t^2 u + \Delta u = 0$ in dimension 3, subject to the initial data $u(0, x) = u_0, \partial_t u(0, x) = 0$ we find,

$$u(t, x) = (2\pi)^{-3} \int_{\mathbb{R}^3} e^{ix\xi} \cos(t|\xi|) \hat{u}_0(\xi) d\xi \tag{87}$$

It pays to make a comparison between the Fourier based formula (87) and the Kirchoff formula (see companion article), which involves only the physical variables (t, x) ,

$$u(t, x) = \partial_t \left((4\pi t)^{-1} \int_{|x-y|=t} u_0(y) da(y) \right) \tag{88}$$

with da the area element of the sphere of radius t centered at $x, |y - x| = t$. Using the Plancherel formula is very easy to deduce from (87) the L^2 bound,

$$\int_{\mathbb{R}^3} |u(t, x)|^2 dx \leq C \|u_0\|_{L^2(\mathbb{R}^3)}^2$$

while obtaining such a bound from (88) seems implausible, since the formula involves a derivative. On the other hand (88) is perfect for giving us domain of influence information. Indeed we read immediately from the formula that if u_0 is supported in ball $B_a = \{|x - x_0| \leq a\}$ than $u(t, x)$ is supported in the ball $B_{a+|t|}$ for any time t . This fact, on the other hand, does not at all seem transparent in the Fourier based formula (87). The fact that different representations of solutions have different, even opposite, strengths and weaknesses has important consequences for constructing parametrices, i.e. approximate solutions, for more complicated, linear variable coefficient or nonlinear wave equations. There are two type of possible constructions, those in physical space, which mimic the physical space formula (88) or those in Fourier space, which mimic formula (87). The first are called Kirchoff-Sobolev, or Hadamard parametrices while the second are called Lax parametrices, or, more generally, Fourier integral operators. In my own research I often had to agonize about which of the two is most suited for the problem at hand.

3) *A priori estimates:* A priori estimates allow us to derive crucial, qualitative, information about solutions to complicated equations without having to solve the equations explicitly. The best known examples are *energy estimates, maximum*

principle or *monotonicity* type arguments. The simplest example of the former type is the following identity, for the Laplace operator $\Delta = \partial_1^2 + \dots + \partial_d^2$ in \mathbb{R}^d ,

$$\|\partial^2 u\|_{L^2(\mathbb{R}^d)}^2 =: \sum_{i,j} \|\partial_i \partial_j u\|_{L^2(\mathbb{R}^d)}^2 = \|\Delta u\|_{L^2(\mathbb{R}^d)}^2 \quad (89)$$

which holds for all functions $u = u(x)$, verifying mild assumptions of decay as $|x| \rightarrow \infty$. From (89) we infer that the equation $\Delta u = f$, with $f \in L^2$, must have bounded second derivatives in $L^2(\mathbb{R}^d)$. The identity can be easily derived by integrating by parts in the integral $\int_{\mathbb{R}^d} (\Delta u)^2 dx$. To see how this is done take for simplicity $d = 2$. Then,

$$\int_{\mathbb{R}^2} \partial_1 \partial_2 u \cdot \partial_1 \partial_2 u dx = - \int_{\mathbb{R}^2} \partial_2 u \cdot \partial_2 \partial_1^2 u dx = \int_{\mathbb{R}^2} \partial_2^2 u \cdot \partial_1^2 u dx$$

Thus,

$$\begin{aligned} \int_{\mathbb{R}^2} |\Delta u|^2 dx &= \int_{\mathbb{R}^2} |\partial_1^2 u + \partial_2^2 u|^2 dx = \int_{\mathbb{R}^2} (|\partial_1^2 u|^2 + |\partial_2^2 u|^2 + 2\partial_1^2 u \cdot \partial_2^2 u) dx \\ &= \int_{\mathbb{R}^2} (|\partial_1^2 u|^2 + |\partial_2^2 u|^2 + |\partial_1 \partial_2 u|^2 + |\partial_2 \partial_1 u|^2) dx \end{aligned}$$

from which (89) follows.

A similar identity holds for the Laplace -Beltrami operator Δ_S on a compact two dimensional Riemannian manifold S , see (18), which we have encountered in section 1. The identity, which can be proved also by the same kind of integration by parts, takes the form,

$$\int_S |\partial^2 u|_g^2 da_g + \int_S K |\partial u(x)|_g^2 da_g = \int_S |\Delta_S u|^2 da_g \quad (90)$$

where K is the Gauss curvature and $|\partial^2 u|_g^2$, $|\partial u|_g^2$ is an expression analogous to $\sum_{i,j} |\partial_i \partial_j u|^2$ and, respectively, $\sum_i |\partial_i u|^2$. Thus, if the Gauss curvature is positive as in the case of the standard sphere, we derive for any solution of the equation $\Delta_g u = f$ the very useful, and painless, a-priori estimate,

$$\|\partial^2 u\|_{L^2(S)} \leq \|f\|_{L^2(S)}$$

Formula (90) is typical to a general, very useful, class of identities in Riemannian geometry called Bochner identities. Similar identities hold also for systems of first order equations, such as Cauchy-Riemann. Consider the vector function $u = (u_1, u_2)$ in \mathbb{R}^2 and the first order operators $\operatorname{div} u = \partial_1 u_1 + \partial_2 u_2$, $\operatorname{curl} u = \partial_2 u_1 - \partial_1 u_2$. Then, if u_1 and u_2 vanish at infinity,

$$\|\partial u\|_{L^2(\mathbb{R}^2)}^2 = \|\operatorname{div} u\|_{L^2(\mathbb{R}^2)}^2 + \|\operatorname{curl} u\|_{L^2(\mathbb{R}^2)}^2 \quad (91)$$

This provides first derivative estimates for the system $\operatorname{div} u = f$, $\operatorname{curl} u = g$ with $f, g \in L^2(\mathbb{R}^2)$. The identity (91) can be generalized to higher dimensional Riemannian manifolds, using the language of differential forms, and is an important ingredient in Hodge theory.

The maximum principle provides another example of a-priori estimate. Consider solutions of the Laplace equation $\Delta u = 0$ on a bounded, connected, domain $D \subset \mathbb{R}^d$ with nice, smooth, boundary ∂D . Assume that u is continuous on the closure of D

and has continuous first and second partial derivatives in the interior of D . Then u must achieve its maximum and minimum values on the boundary, i.e.,

$$\max_{x \in \bar{D}} |u(x)| = \max_{x \in \partial D} |u(x)| \tag{92}$$

Moreover if the maximum, or minimum, is achieved at an interior point of D , u must be constant in D . It is easy to see that u cannot achieve its maximum at interior point if $\Delta u > 0$ in D . Indeed at an interior maximum point all second derivatives of the form $\partial_i^2 u$ must be ≤ 0 which would imply that $\Delta u \leq 0$ at that point. To complete the argument assume $\Delta u = 0$ and consider $v_\epsilon(x) = u(x) + \epsilon|x|^2$ for some $\epsilon > 0$. We have, $\Delta v_\epsilon > 0$ and therefore, v_ϵ must reach its maximum at the boundary ∂D . Therefore, for some constant $M > 0$,

$$\max_{x \in \bar{D}} u(x) \leq \max_{x \in \bar{D}} v_\epsilon(x) \leq \max_{x \in \partial D} v_\epsilon(x) \leq \max_{x \in \partial D} u(x) + \epsilon M$$

Since ϵ can be made arbitrarily small we infer that $\max_{x \in \bar{D}} u(x) = \max_{x \in \partial D} u(x)$. A similar argument can be made for the minimum. The result remains true for solutions to the equation $\Delta u + c(x)u = 0$ provided that $c(x) \leq 0$ in \bar{D} .

The maximum principle provides an important bound for solutions of a PDE with no need to have an explicit representation. Moreover the method is very robust as it can be easily extended to a large class of second order elliptic equations of the form $Lu = 0$, where $L = \sum_{i,j} a^{ij}(x)\partial_i\partial_j + \sum_i b^i(x)\partial_i + c(x)$, with $c \leq 0$.

Another type of a-priori estimates for second order, linear, elliptic PDE, is the Harnack inequality for non negative solutions. In the simplest form it applies to positive solutions of the equation $\Delta u = 0$. More precisely, if u is a non-negative solution of $\Delta u = 0$ in D then, for any domain U whose closure is compact in D , there exists a constant C depending only on U such that,

$$\sup_U u(x) \leq C \inf_U u(x)$$

Like the maximum principle, Harnack inequality can be adapted to second order parabolic equations of the form $\partial_t u = Lu$.

Another important type of a-priori estimates concern extensions of the L^2 bounds of (89) to L^p bounds. The following is the standard example of a Calderon-Zygmund estimate:

$$\sum_{i,j=1}^d \|\partial_i\partial_j u\|_{L^p(\mathbb{R}^d)} \leq C_p \|\Delta u\|_{L^p(\mathbb{R}^d)} \tag{93}$$

The estimate holds for all smooth functions vanishing at infinity and all values $1 < p < \infty$, with a constant C_p depending on p and d . The proof of (93) is far more involved than that of the identity (89) and is based on the theory of singular integral operators. The exceptional cases $p = \infty$ and $p = 1$ are false, yet there are simple modifications of the L^1 and L^∞ norm for which the estimate holds true. In the case of $p = \infty$, the L^∞ norm of a function f can be replaced by the Hölder norm, in a domain $D \subset \mathbb{R}^d$,

$$\|f\|_{C^{0,\alpha}(D)} = \sup_{x,y \in D, x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\alpha}$$

One can show that, for all $0 < \alpha < 1$,

$$\sum_{i,j=1}^d \|\partial_i \partial_j u\|_{C^{0,\alpha}(\mathbb{R}^d)} \leq C_\alpha \|\Delta u\|_{C^{0,\alpha}(\mathbb{R}^d)} \quad (94)$$

which is called a Schauder estimate.

Unlike (89) both estimates (93) and (94) require an explicit representation formula of u in terms of Δu , based on the fundamental solution of Δ . In that sense they are not exactly a-priori. They can however be extended to general second order elliptic equations of the form $Lu(x) = f(x)$ in a domain $D \subset \mathbb{R}^d$ where $Lu(x) = Au(x) + \sum_i b^i(x) \partial_i u(x) + c(x)u(x)$ and,

$$Au(x) = \sum_{i,j} a^{ij}(x) \partial_i \partial_j u(x) \quad (95)$$

the principal part of L , by the so called method of *freezing coefficients*, without relying on explicit representation (which would be hard to get !). To simplify the discussion consider the reduced equation,

$$Au = f, \quad (96)$$

The method of freezing coefficients is based on the fact that locally, in a neighborhood V of a point $x_0 \in \mathbb{R}^n$, the variable coefficient linear operator Au can be well approximated by the constant coefficient operator $A_0 = \sum_{i,j=1}^d a^{ij}(x_0) \partial_i \partial_j$. More precisely it can be shown that the error term $E(x) = \sum_{i,j} (a^{ij}(x) - a^{ij}(x_0)) \partial_i \partial_j u(x)$ is small relative to the Hölder norm of u and its second derivatives in V , i.e.²⁷,

$$\|E\|_{C^{0,s}(V)} \leq \epsilon (\|\partial^2 u\|_{C^{0,s}(V)} + \|u\|_{C^{0,s}(V)})$$

On the other hand, since A is elliptic the matrix $a^{ij}(x_0)$ is positive definite. We can therefore perform a linear change of coordinates $y = Tx$ so that in the new coordinates we have $A_0 = \Delta$. Thus, in V , we can rewrite (96) in the form $\Delta u = f + E$. We can therefore apply an appropriately localized version²⁸ of (94) and deduce that, in a slightly smaller neighborhood $W \subset V$ of x_0 we can get a bound of the form,

$$\|\partial^2 u\|_{C^{0,\alpha}(W)} \leq C (\|f\|_{C^{0,\alpha}(W)} + \|\partial u\|_{C^{0,\alpha}(V)} + \|u\|_{C^{0,\alpha}(V)}) \quad (97)$$

We thus get a Hölder estimate for the second derivatives of u depending not only on f but also u and ∂u . The dependence on these latter can be eliminated by various refinements. This procedure can be worked out further to derive the following:

Theorem[Schauder estimates] *Assume u is a smooth solution of the elliptic equation $Lu = f$ in a bounded domain $D \subset \mathbb{R}^d$, with $L = \sum_{i,j} a^{ij} \partial_i \partial_j + \sum_i b^i \partial_i + c$ as above, subject to the Dirichlet boundary condition $u = 0$ at the boundary of D . Then,*

$$\|\partial^2 u\|_{C^{0,\alpha}(\bar{D})} + \|\partial u\|_{C^{0,\alpha}(\bar{D})} \leq C (\|f\|_{C^{0,\alpha}(\bar{D})} + \|u\|_{C^0(\bar{D})}) \quad (98)$$

with $C^0(\bar{D})$ the uniform norm in D , i.e. $\|u\|_{C^0(\bar{D})} = \sup_{x \in \bar{D}} |u(x)|$. If the maximum principle holds, i.e. if $c \leq 0$, we can eliminate the term $\|u\|_{C^0(\bar{D})}$ on the right.

²⁷assuming that the coefficients a^{ij} are sufficiently smooth

²⁸This requires multiplying u by a cut-off function supported in V and identically one in a small neighborhood V' of x_0

Remark. So far all our examples of a-priori estimates apply to elliptic equations. Many of these estimates extend also to parabolic equations. Unfortunately, there are a lot fewer examples of a-priori estimates for hyperbolic equations. In fact almost all known a-priori estimates for the latter are variations of the energy method. We shall describe the energy method in more details in the next section, here we give a straightforward derivation of the energy identity for the standard wave equation in \mathbb{R}^{1+3} , $\square u = -\partial_t^2 u + \sum_i \partial_i^2 u = 0$. We multiply the equation by $\partial_t u$ and deduce the formula,

$$\frac{1}{2} \partial_t (|\partial_t u|^2 + \sum_i |\partial_i u|^2) - \sum_i \partial_i (\partial_t u \partial_i u) = 0.$$

Integrating this identity in the region $[0, T] \times \mathbb{R}^3$, while assuming²⁹ that u and u_x vanish as $|x| \rightarrow \infty$, we deduce the standard law of conservation of energy,

$$E(t) = E(0), \quad E(t) = \frac{1}{2} \int_{\mathbb{R}^3} (|\partial_t u|^2 + \sum_i |\partial_i u|^2) dx. \quad (99)$$

Exactly the same procedure can be implemented for more general wave equations³⁰ of the form,

$$-a^{00}(t, x) \partial_t^2 u + a^{ij}(t, x) \partial_i \partial_j u = 0, \quad (100)$$

with $a^{00} > 0$ and a^{ij} positive definite. In this case, however, we don't get an exact conservation law, but rather (under specific assumptions for the coefficients a^{00} and a^{ij}), *energy type estimates*. One can show, for example, that if all the first derivatives of the coefficients a are bounded for all $(t, x) \in [0, T] \times \mathbb{R}^3$, then, for any $t \in [0, T]$,

$$\|\partial u(t)\|_{L^2(\mathbb{R}^3)} \leq C \|\partial u(0)\|_{L^2(\mathbb{R}^3)}. \quad (101)$$

A-priori estimates, such as (101), can be used to prove uniqueness of solutions. Indeed if u_1, u_2 are two solutions of (100) with the same initial conditions at $t = 0$ then $v = u_1 - u_2$ is also a solution with zero initial data. Hence, in view of (101) it must vanish in $[0, T] \times \mathbb{R}^3$.

Remark. So far we have seen how to derive L^2 bounds for the first derivatives of solutions to $\square u = 0$, or (100), in terms of the data. It turns out the energy method has a very important extension which allows us to also get bounds on the uniform norm of solutions. Take the case of the wave equation $\square u = 0$ in \mathbb{R}^d . Observe that the equation commutes with the partial derivatives ∂_i , $i = 1, 2, \dots, d$, i.e. $\square \partial_i u = 0$ where ∂ denotes any of the derivatives ∂_i . Applying the energy identity we infer therefore that

$$\|\partial \partial_i u(t)\|_{L^2(\mathbb{R}^d)}^2 \leq \|\partial \partial_i u(0)\|_{L^2(\mathbb{R}^d)}^2, \quad \forall i = 1, \dots, d.$$

Clearly one can continue in the same manner to derive estimates for any number of derivatives of u , i.e. setting $v^\alpha = \partial_1^{\alpha_1} \partial_1^{\alpha_2} \dots \partial_1^{\alpha_d} u$,

$$\|\partial v^\alpha(t)\|_{L^2(\mathbb{R}^d)} \leq \|\partial v^\alpha(0)\|_{L^2(\mathbb{R}^d)}$$

²⁹It turns out that this condition is not at all necessary.

³⁰Similar estimates can be derived for linear symmetric hyperbolic systems, see (75).

for any multi-index $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{N}^d$. To keep track of these energy estimates one introduces the functional norms, called *Sobolev norms*, for functions f in \mathbb{R}^d ,

$$\|f\|_{H^s(\mathbb{R}^d)} = \left(\sum_{|\alpha| \leq s} \|\partial^\alpha f\|_{L^2(\mathbb{R}^d)}^2 \right)^{1/2} \quad (102)$$

One defines the Hilbert space $H^s(\mathbb{R}^d)$ as being the completion of $C_0^s(\mathbb{R}^d)$ relative to this norm. This latter space is simply the space of s times continuously differentiable functions which vanish outside some compact set of \mathbb{R}^d . Using this notation we infer that, any solution of $\square u = 0$ and any $s \geq 0$,

$$\|\partial u(t)\|_{H^s(\mathbb{R}^d)} \leq \|\partial u(0)\|_{H^s(\mathbb{R}^d)} \quad (103)$$

To derive an uniform bound for ∂u one relies on the following, Sobolev embedding theorem³¹.

Theorem[Sobolev] *If $s > \frac{d}{2}$, there exists a constant $C > 0$ such that for any function in $C_0^s(\mathbb{R}^d)$,*

$$\|f\|_{L^\infty(\mathbb{R}^d)} \leq C \|f\|_{H^s(\mathbb{R}^d)} \quad (104)$$

Together with (103) the above theorem allows us to conclude that, if the initial data of a solution u belongs H^{s+1} , i.e. $u(0) \in H^{s+1}(\mathbb{R}^d)$ and $\partial_t u(0) \in H^s(\mathbb{R}^d)$, for $s > \frac{d}{2}$, then the first derivatives of u remain bounded for all time.

Remark 1. The same method can be used to get uniform bounds, in a time interval $[0, T]$, for solutions to the equation (100). The a-priori L^∞ bound thus obtained play an essential role in proving the local existence and uniqueness theorem for nonlinear hyperbolic equations mentioned in the previous section, see Theorem[LEU-Hyperbolic], such as the Einstein equation in wave coordinates.

Remark 2. A more general type of energy estimates, based on using the symmetries of the linear part of the equations, allows one to also prove global in time, results³² for quasilinear wave equations. The starting point of this generalization is based on the observation that the operator \square commutes with a larger class of first order operators, than the partial derivatives ∂_i used above. For example one can show that if u is a solution of $\square u = 0$, so is Γu where Γ can be any of the first order operators $x_i \partial_j - x_j \partial_i$, $t \partial_i + x_i \partial_t$ or $t \partial_t + \sum_i x^i \partial_i$ as well as ∂_i . Thus applying the energy inequality to $w = \Gamma_1 \Gamma_2 \dots \Gamma_k u$, with $\Gamma_1, \Gamma_2, \dots, \Gamma_k$ any of the operators listed above, we derive,

$$\|\partial \Gamma_1 \Gamma_2 \dots \Gamma_k u(t)\|_{L^2(\mathbb{R}^d)}^2 \leq \|\partial \Gamma_1 \Gamma_2 \dots \Gamma_k u(0)\|_{L^2(\mathbb{R}^d)}^2$$

A global version of the Sobolev inequality (104) can then be used to establish not only a uniform bound for ∂u but also a decay estimate. More precisely one can show that, if the initial data for u , and a finite number of their derivatives, decay

³¹As a consequence of the theorem we infer that the completed space $H^s(\mathbb{R}^s)$ consists of bounded continuous functions, i.e. $H^s(\mathbb{R}^s)$ embeds continuously in the space bounded continuous functions endowed with the uniform norm.

³²such as the global stability of the Minkowski space in General Relativity

sufficiently fast at infinity then the corresponding solution for $\square u = 0$ verify an estimate of the form,

$$|\partial u(t, x)| \leq C(1 + |t| + |x|)^{-\frac{d-1}{2}} (1 + ||t| - |x||)^{-\frac{1}{2}}, \quad \forall (t, x) \in \mathbb{R}^{1+d}. \quad (105)$$

In particular, $\max_{x \in \mathbb{R}^d} |\partial u(t, x)| \leq C|t|^{-\frac{d-1}{2}}$, for $|t| \geq 1$. which shows that even though their total energy remains constant in time, solutions of the wave equation decay in time, uniformly in $x \in \mathbb{R}^d$. The inequality (105) gives the additional information that solutions decay faster away from the light cones $t = t_0 \pm |x|$. This corresponds to the physical fact that most of the energy of an electromagnetic wave is carried along characteristic directions. The method of proving (105) is quite robust. It can be used to prove global existence results for general quasilinear wave equations of the type (73), for sufficiently small initial conditions. A variation of the method is used to prove global existence for solutions to the Einstein equations, which are sufficiently close to the Minkowski space³³. We shall mention the result later, it is called the *global stability of Minkowski space*.

There are many other type of important a-priori estimates which appear in the modern theory of nonlinear hyperbolic and dispersive equations. Without giving any details I ought to mention Strichartz and bilinear estimates, which have played an essential role in recent years. Another type of a-priori estimates, Carleman type, appear in connection to ill posed problems. Finally, a-priori estimates can also be used to establish the breakdown of regularity or *blow-up* of solutions to some nonlinear equations. But maybe, to summarize, it suffice to say, without much exaggeration, that a-priori estimates play a fundamental role in just about any aspect of the modern theory of PDE.

4. Boot-strap and continuity arguments: The *boot-strap* argument is a method, or rather a powerful general philosophy³⁴, to derive a priori estimates for nonlinear equations. According to it we start by making assumptions about the solutions we are looking for. These allow us to think of the original nonlinear problem as a linear one whose coefficients satisfy properties consistent with the assumptions. We may then use linear methods, based on a priori estimates, or explicit representation of solutions, to try to show that the solutions to this linear problem behave as well, in fact better, than we have postulated. Here is a very simple theorem which can be proved by this principle.

Theorem. *Consider the second order ODE,*

$$\partial_t^2 u = -V'(u), \quad u(0) = u_0, \quad \partial_t u(0) = u_1 \quad (106)$$

with $V : \mathbb{R} \rightarrow \mathbb{R}$ a smooth function verifying $V(0) = V'(0)$ and $V''(0) > 0$, such as $V(u) = \frac{1}{2}c^2 u^2 - u^3$, for some positive constant $1 \geq c > 0$. Then, for all u_0, u_1 sufficiently small, there exists a unique global solution of the equation, which remains close the origin, i.e. $|u(t)| + |\partial_t u(t)|$ stays small for all $t \geq 0$.

³³Minkowski space is a trivial solution of the Einstein vacuum equations

³⁴One can best characterize this powerful method, which allows us to use linear theory without actually having to linearize the equation, as a *conceptual linearization*.

Since $u = 0$ is an exact solution, corresponding to the initial data $u(0) = 0$, we may expect that solutions, which start close to zero, may remain small for all $t \geq 0$. Indeed, since $|V(u)| \leq C|u|^2$ for small values u we may hope that the nonlinearity does not create any problems. This, however is not necessarily true, as the example $u'(t) = u^2$ demonstrates. Indeed solutions to this equation develop singularities in finite time no matter how small the initial data is, and despite the smallness of u^2 near the origin. The problem is that the linear equation $\partial_t u = 0$, like $\partial_t^2 u = 0$, is unstable relative to small perturbations. What makes the theorem work is that the relevant linearized problem is not $\partial_t^2 u = 0$ but rather the linear oscillator $\partial_t^2 u + c^2 u = 0$. Indeed, using the specific example for V we write our nonlinear equation in the form, $\partial_t^2 u + c^2 u = u^3$. Multiplying both sides by $\partial_t u$ we derive the conservation law,

$$E(t) = E(0), \quad E(t) = \frac{1}{2}((\partial_t u(t))^2 + c^2 u(t)^2) - \frac{1}{3}u(t)^3$$

From this we derive the following conditional a-priori estimate:

Lemma. *As long as $|u(t)| \leq \epsilon < \frac{3}{4}c^2$, we must have $E(0) \geq \frac{c^2}{4}(\partial_t u(t))^2 + u(t)^2$.*

We now start with initial conditions u_0, u_1 such that $E(0) < \delta$, with δ sufficiently small. According to the LEU theorem for ODE, see section 2, there exists a unique classical solution $u \in C^2$, with u and $\partial_t u$ continuous, in a time interval $(-T_*, T_*)$. We make the following *boot-strap assumption*, for a small parameter ϵ to be fixed later,

$$A(T): \quad \sup_{t \in [0, T]} (\partial_t u(t))^2 + u(t)^2 \leq \epsilon^2,$$

Clearly $A(0)$ is verified if δ is sufficiently small relative to ϵ . Let $T_m \leq T_*$ be the largest time for which $A(T)$ holds. By the continuity of u and $\partial_t u$, if T_m is finite, we must have,

$$\sup_{t \in [0, T_m]} (|\partial_t u(t)|^2 + |u(t)|^2) = \epsilon^2. \quad (107)$$

We show that, by choosing ϵ, δ , that in fact, $\sup_{t \in [0, T_m]} |\partial_t u(t)|^2 + u(t)^2 \leq \frac{1}{2}\epsilon^2$. This contradiction will imply that both T_* and T_m must be infinite. We first choose $\epsilon < \frac{3}{4}c^2$ so that, according to our lemma, (107) implies

$$\frac{c^2}{4}(\partial_t u(t))^2 + u(t)^2 \leq E(0) \leq \delta, \quad \forall t \in [0, T_m]$$

Now we can choose δ sufficiently small, so that $\frac{4\delta}{c^2} \leq \frac{1}{2}\epsilon^2$, to infer that $(\partial_t u(t))^2 + u(t)^2 \leq \frac{1}{2}\epsilon^2$ as desired.

Remark. The boot-strap method is indispensable in the modern theory of nonlinear evolution equation. The proof of the nonlinear stability of Minkowski space, mentioned earlier, is based on a huge boot-strap procedure inside which one has to make a number of smaller boot-straps in order to advance the argument.

The above application of the boot-strap principle is typical to evolution problems, where there is a natural time parameter. A related method can be applied to

elliptic problems; it is called the method of continuity. Here is a simple example to illustrate how this works. Consider the semilinear elliptic problem in a domain $D \subset \mathbb{R}^3$,

$$-\Delta u + u^3 = f(x), \quad u|_{\partial D} = 0 \tag{108}$$

The idea of a continuity argument is to start with a linear problem which we know how to solve, such as

$$-\Delta u = f(x), \quad u|_{\partial D} = 0$$

and deform to (108). We thus introduce the family of problems, with $t \in [0, 1]$,

$$-\Delta u + tu^3 = f(x), \quad u|_{\partial D} = 0 \tag{109}$$

If we can show that the set $J \subset [0, 1]$ of values of t for which the problem can be solved in some functional space X , is both open and closed we would infer that $J = [0, 1]$ and therefore (108) can be solved in X . It is not too difficult, using the implicit function theorem and the a-priori estimates (98), to show that J must be open. To do this it would be natural to choose X the space of twice continuously differentiable functions $C^2(\bar{D})$ and take f to belong to the space of continuous functions $Y = C^0(\bar{D})$. But, as we have noticed earlier, these spaces are not well behaved, i.e. if $\Delta u \in C^0(\bar{D})$ it does not follow that $u \in C^2(\bar{D})$. The correct modification is to take $Y = C^\alpha(\bar{D})$ and $X = C^{2,\alpha}(\bar{D})$, with norm,

$$\|u\|_{C^{2,\alpha}(\bar{D})} = \|\partial^2 u\|_{C^{0,\alpha}(\bar{D})} + \|\partial u\|_{C^{0,\alpha}(\bar{D})} + \|u\|_{C^{0,\alpha}(\bar{D})}$$

To show that J is closed one needs a non-linear a-priori estimate for (109) for all values of t . In fact we need to show that, for fixed $f \in C^{0,\alpha}(D)$, any solution u_t of (109) lies in a compact subspace of X . Thus if t_* belongs to the closure of the set J in $[0, 1]$ there must exist a sequence $t_k \in J$ with $t_k \rightarrow t_*$ and functions u_k which solve (109) for $t = t_k$. By compactness we would be able to subtract a subsequence of u_k converging to a function $u \in X$ which solves (109) for $t = t_*$. This is, typically, the more demanding part of the argument. In the case of our equation (109), a first and crucial step, in establishing the needed a-priori estimates is to observe that the maximum principle for the nonlinear operator $\Delta u - tu^3$ holds true. Indeed we can write it in the form $\Delta u + c(x)u$ with $c(x) = -tu^2(x) \leq 0$.

5) *Method of generalized solutions:* The best way to introduce and illustrate the importance of the generalized solutions in PDE is through the *Dirichlet Principle*. This originates in the observation that harmonic functions (i.e. solutions of $\Delta u = 0$) in a bounded domain $D \subset \mathbb{R}^d$ with prescribed Dirichlet boundary condition $u|_{\partial D} = f$ are minimizers of the Dirichlet integral, or Dirichlet functional,

$$\|v\|_{D_r}^2 = \frac{1}{2} \int_D |\nabla v|^2 = \frac{1}{2} \sum_{i=1}^d \int_D |\partial_i v|^2, \tag{110}$$

among all functions (in an appropriate functional space X) which take the specified value f on ∂D . It was Riemann who first had the idea that the Dirichlet principle could be used to solve the Dirichlet problem of finding solutions of $\Delta u = 0$ in a domain D with prescribed values at the boundary,

$$\Delta u = 0, \quad u|_{\partial D} = u_0, \tag{111}$$

by actually minimizing the Dirichlet integral. It seems most natural to take $X = C^2(\bar{D})$, the space of twice continuously differentiable functions on \bar{D} with norm, $\|v\|_{C^2(\bar{D})} = \sup_{x \in D} (|v(x)| + |\partial v(x)| + |\partial^2 v(x)|)$. Another reasonable choice is to take $X = C^1(\bar{D})$ with norm, $\|v\|_{C^1(\bar{D})} = \sup_{x \in D} (|v(x)| + |\partial v(x)|)$. Indeed, observe that the Dirichlet norm $\|v\|_{D_r}$ is finite for $v \in C^1(\bar{D})$. The precise formulation of the Dirichlet principle, which requires the definition of the Sobolev spaces $H^1(D)$ and $H_0^1(D)$ as the correct function spaces in which the minimization takes place, has a fascinating history starting with the bold but flawed attempt by Riemann, who did in fact chose $X = C^2(\bar{D})$, followed by a penetrating criticism of Weierstrass, who showed that the functional may not achieve its minimum in either $C^2(\bar{D})$ or $C^1(\bar{D})$ space, and then the revival and final triumph of the principle through a long, inspiring, process of defining appropriate *functional spaces*, introduction of *generalized solutions* (using Hilbert space methods) and the development of a *regularity theory* for these. The fundamental concept of generalized solutions was introduced in order to circumvent the flaw in the original argument of Riemann, due to the incompleteness of the spaces $C^1(\bar{D}), C^2(\bar{D})$ relative to the Dirichlet norm $\|\cdot\|_{D_r}$, see (110). It was first observed by B. Levi that a sequence of functions which minimizes the Dirichlet integral is a Cauchy sequence in the normed space $X = (C^1(\bar{D}), \|\cdot\|_{D_r})$ and thus converges to a function u , in an appropriate completion of X . This is precisely the space called $H^1(D)$. The resulting limiting function u fails, however, to be in $C^2(D)$ and thus we cannot check whether $\Delta u = 0$ in D . Nevertheless we can interpret u as a *generalized solution* of the equation in the following sense:

$$\sum_i \int_D \partial_i u \partial_i \phi = 0, \quad (112)$$

for any smooth function ϕ which vanishes at the boundary ∂D . These are called test functions. Observe that this is a reasonable generalization of a classical solution. Indeed, if u is actually a classical, twice continuously differentiable solution, i.e. $u \in C^2(D)$, then we can integrate by parts in (112) and derive $\int \Delta u \phi = 0$ from which, since ϕ is an arbitrary test functions, we infer $\Delta u = 0$.

We thus have a generalized solution of our equation, yet this is not exactly what we were looking for. In fact it is not even clear in what sense this generalized solution satisfy our boundary condition $u|_{\partial D} = u_0$. To deal properly with the boundary condition it pays to modify the above procedure a little. The idea is to solve, instead³⁵ of (111) the inhomogeneous equation

$$\Delta u = f, \quad u|_{\partial D} = 0 \quad (113)$$

with zero boundary condition. Then, instead of the Dirichlet integral, one minimizes the modified functional,

$$I(v) = \|v\|_{D_r}^2 - \int_D v(x) f(x) dx$$

relative to the completion of the set $C_0^\infty(D)$ of smooth functions with compact support in D , with respect to the Dirichlet norm. One obtains this way the Sobolev space $H_0^1(D)$. The minimization produces, by a simple Hilbert space argument³⁶,

³⁵It is not hard to show that the two problems are equivalent.

³⁶It follows from the Riesz representation theorem.

a weak solution of $\Delta u = f$, i.e. a function u which verifies,

$$-\sum_i \int_D \partial_i u \partial_i \phi = \int_D f \phi. \quad (114)$$

It remains to show that if the domain D is sufficiently nice and f is sufficiently smooth (at least continuous), our weak solution is in fact a classical solution in $C^2(D) \cap C^0(\bar{D})$. This is the hard part of the proof and is based on an extension of regularity theory of elliptic equations to generalized solutions.

To summarize, the method outlined above rests on two steps. In the first step one produces a generalized solution of a linear, or nonlinear PDE, based on a minimization procedure. The task in the second step is to show that the generalized solution is in fact a classical solution. In some situations, however, the generalized solution may turn out to have singularities, in which case the challenge becomes to understand their nature and prove realistic partial regularity results, i.e. show that the generalized solution is smooth every where except a small exceptional set.

The minimization procedure is very natural for equations which can be derived from a variational principle³⁷, but other methods for producing generalized solutions can be devised. It is important however that the original equation has a divergence structure such as the equation

$$Lu = \sum_{i,j=1}^d \partial_i (a^{ij}(x) \partial_j u) + Mu = 0 \quad (115)$$

with M the first order operator, $Mu(x) = \sum_i b^i(x) \partial_i u(x) + c(x)u(x)$. This structure is always present for equations derived by a variational principle.

Definition. We say that u is a formal weak solution of (115) in a domain $D \subset \mathbb{R}^d$ if, for any test function $\phi \in C_0^1(D)$, i.e. continuously differentiable and compactly supported in D ,

$$-\sum_{i,j=1}^d \int_D a^{ij}(x) \partial_j u(x) \partial_i \phi(x) + \int_D Mu(x) \phi(x) = 0 \quad (116)$$

Observe that the integral (116) makes sense, merely if the coefficients a^{ij} , b^i and c are measurable functions, bounded almost everywhere on D , and both u and its first derivatives are just integrable in D . In particular we can make sense of a generalized solution for $u \in H^1(D)$, the Sobolev space introduced above.

A typical example of the second step is embodied in the De Giorgi-Nash technique which allows one to derive full regularity estimates for the generalized solutions of elliptic equations of the form (115). More precisely we have the following fundamental result.

Theorem *Assume that a^{ij}, b, c are measurable, bounded almost everywhere (a.e.) in D and that a^{ij} verify the ellipticity condition,*

$$a^{ij}(x) \xi_i \xi_j > c |\xi|^2, \quad \forall \xi \in \mathbb{R}^d, \quad \text{and a.e. } x \in D$$

³⁷see definition next section

Then, every generalized solution $u \in H^1(D)$ of the equation $Lu = 0$, i.e. for which (116) holds, must be continuous in D , and in fact Hölder continuous for some exponent $\delta > 0$, i.e. $u \in C^{0,\delta}(D)$.

This theorem has wide range of applications to nonlinear elliptic problems. It provides, in particular the solution to the famous problem³⁸ of the regularity of minimal hypersurfaces, as graphs over convex domains, in all dimensions ≤ 7 . Other important applications of the Nash -De Giorgi method were found in connection to the Calabi problem in Complex Geometry, Ricci flow, free boundary value problem in Continuum Mechanics etc. It is important to remark, however, that the De Giorgi-Nash result does not extend to systems of second order equations in divergence form.

Though generalized solutions are most effective for elliptic problems, their range of applicability encompasses all PDE. The idea of a generalized solution appears in fact already in the work of D'Alembert in connection to the one dimensional wave equation(vibrating string). As we have seen the fundamental solutions to the basic linear equations have to be interpreted as distributions, which are examples of generalized solutions. As we have seen a systematic and compelling concept of generalized solutions has developed in connection to the Dirichlet Principle and, more generally, for other variational problems via what is called *the direct variational method*.

The notion of generalized solutions has also proved successful to nonlinear evolution problems, such as systems of conservation laws in one space dimension. An excellent example is provided by the Burger equation (39). As we have seen solutions to $\partial_t u + u\partial_x u = 0$ develop singularities in finite time no matter how smooth the initial conditions are. It is natural to ask whether solutions continue to make sense, as generalized solutions, even beyond the time when these singularities form. A natural notion of generalized solution is the following,

$$\int_{\mathbb{R}^{1+1}} u\partial_t\phi + \frac{1}{2} \int_{\mathbb{R}^{1+1}} u^2\partial_x\phi = 0, \quad \forall\phi \in C_0^\infty(\mathbb{R}^{1+1})$$

It can be shown that, under additional conditions, called entropy conditions, the IVP for the Burger equation admits a unique, global, i.e. for all $t \in \mathbb{R}$, generalized solution. Today we have a satisfactory theory of global solutions, with small bounded variation, to a large class of hyperbolic systems of one dimensional conservation laws, called *strictly hyperbolic*.

The question of what is a good concept of a generalized solution for more complicated nonlinear evolution equations, though fundamental, is far more murky. For higher dimensional evolution equations the first concept of a *weak solution* was introduced by J. Leray. I call weak a generalized solution for which one cannot prove any type of uniqueness. This unsatisfactory situation may be temporary, due to our technical inabilities, or unavoidable in the sense that the concept itself is flawed. Leray was able to produce, by a compactness method, a weak solution of the initial value problem for the Navier-Stokes equations. The great advantage

³⁸This problem is a straightforward extension to higher dimensions of the problem of finding minimal surfaces, i.e. solutions to (6), with a prescribed Dirichlet condition at the boundary.

of the compactness method (and its modern extensions which can, in some cases, cleverly circumvent lack of compactness) is that it produces global solutions for all data. This is particularly important for supercritical, or critical, nonlinear evolution equations³⁹ where we expect that classical solutions develop finite time singularities. The problem, however, is that one has very little control of these solutions, in particular we don't know how to prove their uniqueness⁴⁰. Similar type of solutions were later introduced for other important nonlinear evolution equations. In most of the interesting cases of supercritical evolution equations, such as Navier-Stokes, the usefulness of the type of weak solutions used so far remains undecided.

Finally, I need to remark that generalized solution can be defined even for nonlinear equations which are not in conservation form, such as the Hamilton-Jacobi equations (42). The generalized solutions introduced in this context are called *viscosity solutions*. Recall that classical solutions of the Hamilton-Jacobi equations become singular in finite time. The viscosity solutions provide the correct notion of generalized solution beyond singularities.

6) *Micro-local analysis, Parametrics and paradifferential calculus:* One of the fundamental difficulties of hyperbolic and dispersive equations consists of the interplay between geometric properties, which concern the physical space, and properties, intimately tied to oscillations, which are best seen in Fourier space. Micro-local analysis is a general, still developing, philosophy according to which one isolates the main difficulties by careful localizations in physical or Fourier space, or in both. An important application of this point of view is the construction of parametrics for linear hyperbolic equations and their use in propagation of singularities results. Parametrics, as we have already mentioned, are approximate solutions of linear equations with variable coefficients, modulo error terms which are smoother. The paradifferential calculus is an extension of the micro-local analysis to nonlinear equations. It allows one to manipulate the form of a nonlinear equation, by taking account of the way large and small frequencies interact, to achieve remarkable technical versatility.

7) *Scaling properties of nonlinear equations.* Essentially all basic nonlinear equations have well defined scaling properties. Take for example, the Burger equation (39), $\partial_t u + u \partial_x u = 0$ and observe that if u is a solution of the equation so is $u_\lambda(t, x) = u(\lambda t, \lambda x)$. Similarly, if u is a solution of the cubic nonlinear Schrödinger equation in \mathbb{R}^d ,

$$i \partial_t u + \Delta u + c |u|^2 u = 0,$$

so is $u_\lambda(t, x) = \lambda u(\lambda^2 t, \lambda x)$. The relationship between the nonlinear scaling of the equation and the a-priori estimates available for solutions to the equations leads to an extremely useful classification between sub-critical, critical and supercritical equations which will be discussed in more details in section 5. For the moment it suffices to say that subcritical equations are those for which the nonlinearity can be controlled by the existing a-priori estimates of the equation, while supercritical

³⁹We discuss these notions later in the article.

⁴⁰Leray was very concerned about this point. Though, like all other researchers after him, he was unable to prove uniqueness of his weak solution, he showed however that it must coincide with a classical one as long as the latter does not develop singularities.

are those for which the nonlinearity appears to be stronger. Critical equations are borderline. The definition of criticality and its connection to the issue of regularity, plays a very important heuristic role in nonlinear PDE. One expects that supercritical equations develop singularities while subcritical equations don't. We will talk more about this issue in section 5.

4. MAIN EQUATIONS

In the last section we have pointed out that, while there is no hope to find a general theory of all PDE, there exists nevertheless a wealth of general ideas and techniques whose knowledge is relevant in the study of all, or most, important equations. In this section we indicate how it may be possible to identify what features characterize the equations we call important.

Most of our basic PDE can be derived from simple geometric principles, which happen to coincide with some of the underlying geometric principles of modern physics. These simple principles provide a unifying framework⁴¹ for our subject and help endow it with a sense of purpose and cohesion. They also explain why a very small number of linear differential operators, such as the Laplacian and D'Alembertian, are all pervasive; they are the simplest approximations to equations naturally tied to the two most fundamental geometric structures, Euclidean and Minkowskian. The Heat equation is the simplest paradigm for diffusive phenomena while the Schrödinger equation can be viewed as the Newtonian limit of the Klein Gordon equation. The geometric framework of the former is Galilean space which, itself, is simply the Newtonian limit of the Minkowski space⁴².

Starting with the Euclidean space \mathbb{R}^n , the Laplacean Δ is the simplest differential operator invariant under the group of isometries, or rigid transformations, of \mathbb{R}^n . The heat, Schrödinger, and wave operators $\partial_t - \Delta$, $\frac{1}{i}\partial_t - \Delta$ and $\partial_t^2 - \Delta$ are the simplest evolution operators which we can form using Δ . The wave operator $\square = -\partial_t^2 + \Delta$ has a deeper meaning, however, it is associated to the Minkowski space \mathbb{R}^{n+1} in the same way that Δ is associated to \mathbb{R}^n . Moreover, the solutions to $\Delta\phi = 0$ can be viewed as special, time independent solutions, to $\square\phi = 0$. The Schrödinger equation can also be obtained, by a simple limiting procedure, from the Klein-Gordon operator $-c^{-2}\partial_t^2 + \Delta - m^2c^2$ by letting c , the velocity of light, tend to ∞ . Appropriate, invariant, and local definitions of square roots of Δ and \square , or $\square - k^2$, corresponding to spinorial representations of the Lorentz group, lead to the associated Dirac operators, see (16). In the same vein we can associate to every Riemannian, or Lorentzian, manifold (M, g) the operators Δ_g , resp \square_g , or the corresponding Dirac operators. These equations inherit in a straightforward way the symmetries of the spaces on which they are defined.

⁴¹The scheme presented below is only an attempt to show that, in spite of the enormous number of PDE's studied by mathematicians, physicists and engineers, there are nevertheless simple basic principles which unite them. I don't want, by any means, to imply that the equations discussed below are the only ones worthy of our attention.

⁴²This is done by starting with the Minkowski metric $m = \text{diag}(-1/c^2, 1, 1, 1)$, where c corresponds to the velocity of light, and letting $c \rightarrow \infty$.

4.1. Variational equations. There exists a general, extremely effective, scheme of generating equations with prescribed symmetries, which we describe below. One starts with a scalar quantity, called *Lagrangian*, such as,

$$\mathcal{L}[\phi] = \sum_{\mu, \nu=0}^3 m^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \quad (117)$$

with $\phi : \mathbb{R}^{1+3} \rightarrow \mathbb{R}$ and V is a real function of ϕ such as, for example, $V(\phi) = \phi^3$. Here ∂_μ denote the partial derivatives with respect to the coordinates x^μ , $\mu = 0, 1, 2, 3$ and $m^{\mu\nu} = m_{\mu\nu}$ denotes the Minkowski metric, given by the diagonal matrix, $\text{diag}(-1, 1, 1, 1)$. In what follows we shall use the summation convention over repeated indices. Thus instead of $\sum_{\mu, \nu=0}^3 m^{\mu\nu} \partial_\mu \phi \partial_\nu \phi$ we simply write $m^{\mu\nu} \partial_\mu \phi \partial_\nu \phi$. We also raise, and lower, indices with respect to the metric m . More precisely for a one covector A_μ we write $A^\mu = m^{\mu\nu} A_\nu$. Similarly, if B^μ is a vector, we define $B_\mu = m_{\mu\nu} B^\nu$. Thus, if we denote $x^0 = t$, then $x_0 = m_{0\nu} x^\nu = -t$.

We associate to $\mathcal{L}[\phi]$, defined above, the so called *action integral*,

$$\mathcal{S}[\phi] = \int_{\mathbb{R}^{3+1}} \mathcal{L}[\phi]. \quad (118)$$

with the integration taken relative to the standard measure of \mathbb{R}^{1+3} . Observe that both $\mathcal{L}[\phi]$ and $\mathcal{S}[\phi]$ are invariant under translations and Lorentz transformations. In other words if $T : \mathbb{R}^{1+3} \rightarrow \mathbb{R}^{1+3}$ is an isometry, i.e. it keeps the metric invariant, and $\phi_T = \phi \circ T$ then $\mathcal{L}[\phi_T] = \mathcal{L}[\phi]_T$ and $\mathcal{S}[\phi_T] = \mathcal{S}[\phi]$.

One defines a compact variation of ϕ to be a smooth one-parameter family of functions $\phi^{(s)} : \mathbb{R}^{1+3} \rightarrow \mathbb{R}$, defined for $s \in (-\epsilon, \epsilon)$, such that, at $s = 0$ we have $\phi^{(0)} = \phi$ and, at all points p outside a compact set of \mathbb{R}^{1+3} , we have $\phi^{(s)} = \phi$. Given such a variation we denote,

$$\delta\phi := \dot{\phi} := \left. \frac{d\phi^{(s)}}{ds} \right|_{s=0}.$$

Thus, for small s , $\phi^{(s)} = \phi + s\dot{\phi} + O(s^2)$, where $O(s^2)$ denote an error term which can be bounded, in absolute value, by a positive constant C multiplied by s^2 , for sufficiently small values of s .

Definition. A field⁴³ ϕ is said to be *stationary* with respect to \mathcal{S} if, for any compact variation $\phi^{(s)}$ of ϕ , we have

$$\left. \frac{d}{ds} \mathcal{S}[\phi^{(s)}] \right|_{s=0} = 0.$$

Variational Principle. *The variational principle, or principle of least action states that an acceptable solution of a given physical system must be stationary with respect to the action integral associated to the Lagrangian of the system. The principle allows us to associate to the given lagrangian, a system of partial differential equations, called the Euler-Lagrange equations.*

⁴³The definition applies to more general fields. In this case a field refers to a function ϕ .

We illustrate the variational principle below by showing that the nonlinear wave equation in \mathbb{R}^{1+3} ,

$$\square\phi - V'(\phi) = 0, \quad (119)$$

is the Euler-Lagrange equation associated to the Lagrangean (117). Given a compact variation $\phi^{(s)}$ of ϕ , we set $\mathcal{S}(s) = \mathcal{S}[\phi^{(s)}]$. Integration by parts gives,

$$\begin{aligned} \frac{d}{ds}\mathcal{S}(s) \Big|_{s=0} &= \int_{\mathbb{R}^{3+1}} [-m^{\mu\nu}\partial_\mu\dot{\phi}\partial_\nu\phi - V'(\phi)\dot{\phi}] \\ &= \int_{\mathbb{R}^{3+1}} \dot{\phi}[\square\phi - V'(\phi)]. \end{aligned}$$

In view of the action principle and the arbitrariness of $\dot{\phi}$ we infer that ϕ must satisfy equation (119). Thus (119) is indeed the Euler Lagrange equation associated to the Lagrangean $\mathcal{L}[\phi] = m^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)$.

Next we show that the Maxwell equations of electromagnetism are also variational, i.e. they can be derived from a Lagrangean. We start with $A = A_\mu dx^\mu$ a one form in Minkowski space \mathbb{R}^{1+3} , and take its exterior derivative $F = dA$. The components of F are simply $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Observe that $dF = 0$ i.e.,

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0. \quad (120)$$

We define the electromagnetic Lagrangean to be the scalar quantity,

$$\mathcal{L}[F] = \frac{1}{2}F_{\mu\nu}F^{\mu\nu} \quad (121)$$

Instead of varying F , however, in the action integral $\mathcal{S}[F] = \int_{\mathbb{R}^{3+1}} \mathcal{L}[F]$, we vary with respect to A instead. In other words we take $F^{(s)} = dA^{(s)}$ where $A^{(s)}$ is a variation of A . Thus $F_{\mu\nu}^{(s)} = \partial_\mu A_\nu^{(s)} - \partial_\nu A_\mu^{(s)}$. Let $\mathcal{S}[s] = \int_{\mathbb{R}^{3+1}} \mathcal{L}[F^{(s)}]$. Differentiating with respect to s and then integrating by parts gives,

$$\frac{d}{ds}\mathcal{S}(s) \Big|_{s=0} = \int_{\mathbb{R}^{3+1}} F^{\mu\nu}(\partial_\mu\dot{A}_\nu - \partial_\nu\dot{A}_\mu) = 2 \int_{\mathbb{R}^{3+1}} (\partial_\nu F^{\mu\nu})\dot{A}_\nu$$

where $\dot{A}_\nu = \frac{d}{ds}A_\nu^{(s)}|_{s=0}$. Therefore, since the variation \dot{A} is arbitrary, we infer that,

$$\partial_\nu F^{\mu\nu} = 0. \quad (122)$$

Together (120) and (122) are the Maxwell equations in space-time form, see compendium article.

A similar derivation holds for the Yang-Mills equations. In this case we take the 1-form A to be valued in a Lie algebra, such as the Lie algebra of 3×3 real antisymmetric matrices. In other words each A_μ is now a 3×3 antisymmetric matrix valued function on \mathbb{R}^{1+3} . In this case one defines $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$ with the bracket denoting the commutator of 3×3 matrices $[A, B] = A \cdot B - B \cdot A$. By a straightforward calculation we find the following equation, analogous to (120),

$$D_\lambda F_{\mu\nu} + D_\mu F_{\nu\lambda} + D_\nu F_{\lambda\mu} = 0, \quad D_\lambda F_{\mu\nu} = \partial_\lambda F_{\mu\nu} + [A_\lambda, F_{\mu\nu}] \quad (123)$$

The Lagrangean of the Yang-Mills equations, is given by,

$$\mathcal{L}[F] = \frac{1}{2}Tr(F_{\mu\nu} \cdot {}^t F^{\mu\nu}) \quad (124)$$

where tF denotes the transposed matrix to F and Tr the usual trace of matrices. Proceeding exactly as in the derivation of (122) we find,

$$D_\nu F^{\mu\nu} = 0. \quad (125)$$

Together (123) and (125) form the Yang-Mills equations, see compendium article.

Another interesting example are wave maps. One starts with maps defined from the Minkowski space \mathbb{R}^{1+d} to a fixed Riemannian manifold M of arbitrary dimension, say 2 for simplicity, with metric h . Take as action integral,

$$\int_{\mathbb{R}^{1+d}} m^{\mu\nu} \langle \partial_\mu \phi, \partial_\nu \phi \rangle_h. \quad (126)$$

Here the derivatives $\partial_\mu \phi$ belong to the tangent space to N and \langle, \rangle_h denotes the Riemannian scalar product on N . The corresponding Euler-Lagrange equation is called the wave maps equations. In local coordinates on N the map ϕ has components ϕ^1, ϕ^2 which verify the equations,

$$\square \phi^I + \Gamma_{JK}^I(\phi) m^{\mu\nu} \partial_\mu \phi^J \partial_\nu \phi^K = 0 \quad (127)$$

The coefficients Γ_{JK}^I are called Christoffel symbols, they are explicit functions depending on the derivatives of the metric coefficients h_{11}, h_{12}, h_{22} , relative to the given coordinate system.

One define in the same manner harmonic maps from euclidean space \mathbb{R}^d to M . In that case we simply replace m with the euclidean metric. The harmonic maps equations take the form,

$$\Delta \phi^I + \Gamma_{JK}^I(\phi) e^{ab} \partial_a \phi^J \partial_b \phi^K = 0$$

with e the euclidean metric. Observe that harmonic maps are time independent solutions to (127). One can naturally extend this construction to harmonic maps between any two Riemannian manifolds.

The Einstein field equations are also variational. In that case the action integral is the integral of the scalar curvature relative to the metric and the variation is done with respect to the metric itself. Other equations of fundamental importance in Hydrodynamics, Continuum Mechanics, or Magneto-hydrodynamics also have a variational structure.

Remark 1. The variational principle only asserts that the acceptable solutions of a given system is stationary; in general we have no reason to expect that the desired solutions minimize or maximize the action integral. This is typically for systems which are time dependent such as the Maxwell, Yang-Mills, Wave Maps and Einstein equations.

There exists however a large class of variational problems, corresponding to time independent physical systems, or geometric problems, for which the desired solutions are in fact extremal. The simplest example is that of geodesics in a Riemannian manifold M which are minimizers⁴⁴ with respect to the length functional. The

⁴⁴This is true, in general, only for sufficiently short geodesics, i.e. passing through two points closed to each other.

length functional associates to any curve γ , passing through two fixed points of M , the corresponding length of the curve $L(\gamma)$, which plays the role of an action integral. In this case a geodesic is not just a stationary point for the functional but a minimum. We have also seen earlier, according to the Dirichlet principle, that solutions to the Dirichlet problem (111) minimize the Dirichlet integral (110). Another example is provided by the minimal surface equation (6) whose solutions are minimizers of the area integral,

$$\int_D (1 + |\partial u|^2)^{\frac{1}{2}} dx.$$

The study of minimizers of various functionals, i.e action integrals, is a venerable subject in mathematics known under the name of Calculus of Variations. A typical problem in the subject is the study of minimizers associated to a general functional,

$$\int_D F(x, u(x), \partial u(x)) \quad (128)$$

where $u : D \subset \mathbb{R}^d \rightarrow \mathbb{R}$, subject to the boundary condition $u|_{\partial D} = u_0$. Here we consider variations $u^{(s)}$ of u such that $u^{(s)}|_{\partial D} = u_0$. Under a convexity assumption for the Lagrangean F , with respect to ∂u , the corresponding Euler-Lagrange equation is an elliptic second order equation in D . It was conjectured by Hilbert in 1900, known as the Hilbert's 19th problem, that, subject to a convexity assumption as well as smoothness of F , regularity of D and of the boundary condition u_0 , the minimizers of (128) always exist and are smooth functions in \bar{D} . Today the problem is well understood, based on the notion of generalized solutions and methods of proving regularity discussed in the previous section. The situation where u in (128) is not a scalar but a vector function the convexity assumption on F is replaced by the important notion of quasi-convexity introduced by Morrey, which allows one to establish the existence of minimizers. However the regularity result may fail in general. This is still an area of active investigation with deep applications to elasticity theory and geometry.

Associated to the variational principle we have another fundamental principle which we state below.

Noether's principle. *To any continuous one parameter group of symmetries of the Lagrangian there corresponds a conservation law for the associated Euler-Lagrange PDE.*

To see how this works out for (117) we consider the following expression, called the energy momentum tensor for the equation (119),

$$T_{\alpha\beta} = \frac{1}{2} \left[\partial_\alpha \phi \partial_\beta \phi - \frac{1}{2} m_{\alpha\beta} (m^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + 2V(\phi)) \right]. \quad (129)$$

Observe that $T_{\mu\nu}$ is symmetric, i.e. $T_{\mu\nu} = T_{\nu\mu}$, and divergenceless, i.e. whenever ϕ verifies (119) we have,

$$\partial^\nu T_{\mu\nu} = 0. \quad (130)$$

For $\mu = 0$, we can write,

$$-\partial_t T_{00} + \sum_{i=1}^3 \partial_i T_{0i} = 0. \quad (131)$$

Integrating (131) on the hyperplane Σ_t , corresponding to a fix value of t , i.e relative to the space variables $x = x^1, x^2, x^3$, we derive the conservation law,

$$E(t) = E(0), \quad \text{where} \quad E(t) = \int_{\Sigma_t} T_{00} \quad (132)$$

The conserved quantity,

$$E(t) = \int_{\Sigma_t} \left(\frac{1}{2} (\partial_t \phi)^2 + \frac{1}{2} \sum_{i=1}^3 (\partial_i \phi)^2 + V(\phi) \right) dx \quad (133)$$

is called the total energy at time t . The identity (133) is called the law of conservation of energy. It clearly coincides with the definition we gave in (99) for the particular case $V = 0$.

The conservation law (133) is the simplest example of a more general procedure which relates symmetries of the underlying space to conservation laws. More precisely consider a vectorfield $X = X^\mu \partial_\mu$. Let $P_\mu = T_{\mu\nu} X^\nu$. By a simple calculation, using both the symmetry and *divergence-less* property (130) of T we derive,

$$\partial^\mu P_\mu = \frac{1}{2} \pi_{\mu\nu} T^{\mu\nu} \quad (134)$$

where $\pi_{\mu\nu} = \partial_\mu X_\nu + \partial_\nu X_\mu$ is the deformation tensor of X .

Definition. *A vectorfield X is said to be Killing if its deformation tensor π vanishes identically.*

For any Killing vectorfield X the right hand side of (134) vanishes. Therefore, as in (131),

$$-\partial_t (T_{\mu 0} X^\mu) + \sum_{i=1}^3 \partial_i (T_{\mu i} X^\mu) = 0$$

Integrating with respect to x^1, x^2, x^3 , we derive the conservation law,

$$\int_{\Sigma_t} T_{\mu 0} X^\mu = \int_{\Sigma_0} T_{\mu 0} X^\mu \quad (135)$$

Observe that (132) is a particular case of (135) corresponding to the Killing vectorfield $X = \partial_t$. Other conservation laws can be derived by choosing $X = \partial_\mu$, $X = t\partial_i + x_i\partial_t$ or $X = x_i\partial_j - x_j\partial_i$, for $i, j = 1, 2, 3$. One can show that these Killing vector-fields are the generators of the one parameter groups of translations, rotations of the spatial coordinates and Lorentz transformations. As we have seen these transformations keep the Lagrangean (117) invariant. This confirms Noëther's principle in the case of equation (119).

The conservation laws corresponding to $X = \partial_i$ and $X = x_i\partial_j - x_j\partial_i$ are called the linear momentum, and, respectively, angular momentum conservation laws. Observe that the quantity $T_{\mu 0} X^\mu$ is positive at all points of Σ_t only in the case

$X = \partial_t$, corresponding to the energy (133). The positivity of the integrand in (133) makes the energy conservation law (132) extremely useful analytically, as it allows us to derive an absolute bound on the size of an arbitrary solution of (119), at all times $t \in \mathbb{R}$, depending only on the size of the initial data at $t = 0$. We say that the a-priori estimate, provided by the energy is *coercive*.

Remark. The method of deriving conserved quantities, described above, can be extended to vectorfields X , called conformal Killing, for which the deformation tensor π is proportional to the metric. Of particular importance is the Morawetz vectorfield $K_0 = (t^2 + |x|^2)\partial_t + 2tx^i\partial_i$.

One can also associate energy momentum quantities, such as (129), to the Maxwell and Yang-Mills equation. The method of deriving conserved quantities from the energy momentum tensor works without any modifications.

4.2. Other equations. Many other familiar equations can be derived from the *variational ones* described above by the following procedures:

(a.) *Symmetry reductions:* Are obtained by assuming that the solutions we are looking for have certain continuous symmetries. They lead to much simpler equations than the original, often intractable, ones. Another, somewhat more general, way of obtaining simpler equations is to look for solutions which verify a certain ansatz, such as stationary⁴⁵, spherically symmetric⁴⁶, self-similar⁴⁷, traveling waves⁴⁸, etc. Typically, the equations obtained by such reductions have themselves a variational structure. In fact the symmetry reduction can be implemented directly on the original lagrangean.

(b.) *The Newtonian approximation and other limits:* We can derive a large class of new equations, from the basic ones described above, by taking one or more characteristic speeds to infinity. The most important one is the Newtonian limit, which is formally obtained by letting the velocity of light go to infinity. As we have mentioned above the Schrödinger equation itself can be derived, in this fashion, from the linear Klein-Gordon equation. In the same way we can formally derive the Lagrangian of non-relativistic Elasticity, Fluids or MHD equations. The formal Newtonian limit of the full Einstein field equations leads to the various continuum mechanics theories in the presence of Newtonian gravity. It is interesting to remark that the non-relativistic equations, look more messy than the relativistic ones. The simple geometric structure of the original equations gets lost in the limit. The remarkable simplicity of the relativistic equations is a powerful example of the importance of Relativity as a unifying principle.

Once we are in the familiar world of Newtonian physics we can perform other well known limits. The famous incompressible Euler equations are obtained by taking the limit of the general non-relativistic fluid equations as the speed of sound

⁴⁵i.e. time independent

⁴⁶i.e. invariant under rotations of the space variables

⁴⁷i.e. solutions which are functions of x/t^a

⁴⁸i.e. functions of $(x - vt)$ for a given velocity vector v .

tends to infinity. Various other limits are obtained relative to other characteristic speeds of the system or in connection with specific boundary conditions, such as the boundary layer approximation in fluids. The equations of Elasticity, for example, approach in the limit, when all characteristic speeds tend to infinity, to the familiar equations of a rigid body in Classical Mechanics. Another important type of limit, leading to a Hamilton-Jacobi equation of the form (42), is the high frequency or the geometric optics approximation.

Many of these very singular limits remain purely formal. While some of them have been rigorously derived, many more present serious analytic difficulties.

(c) *Phenomenological assumptions:* Even after taking various limits and making symmetry reductions, the equations may still remain unyielding. In various applications it makes sense to assume that certain quantities are small and may be neglected. This leads to simplified equations which could be called *phenomenological*⁴⁹ in the sense that they are not derived from first principles. They are used to illustrate and isolate important physical phenomena present in complicated systems. A typical way of generating interesting phenomenological equations, is to try to write down the simplest model equation which describes a particular feature of the original system. Thus, the self-focusing, plane wave effects of compressible fluids, or elasticity, can be illustrated by the simple minded Burgers equation $u_t + uu_x = 0$. Nonlinear dispersive phenomena, typical to fluids, can be illustrated by the famous KdV equation $u_t + uu_x + u_{xxx} = 0$. The nonlinear Schrödinger equation (117) provides a good model problem for nonlinear dispersive effects in Optics. The Ginzburg-Landau equations provide a simple model equation for symmetry breaking, phase transitions. The Maxwell-Vlasov equations is a simplified model for the interactions between electromagnetic forces and charged particles, used in Plasma Physics.

When well chosen, a model equation leads to basic insights into the original equation itself. For this reason simplified model problems are also essential in the day to day work of the rigorous PDE mathematician, who tests ideas on carefully selected model problems. It is crucial to emphasize that good results concerning the basic physical equations are rare; a very large percentage of important rigorous work in PDE deals with simplified equations selected, for technical reasons, to isolate and focus our attention on some specific difficulties present in the basic equations.

It is not at all a surprise that the equations derived by symmetry reductions, various limits and phenomenological assumptions have additional symmetries and therefore additional conservation laws. It is however remarkable that some of them have infinitely many conserved quantities or turn out to be even integrable⁵⁰. The discovery of the integrability of the KdV equation and, later, that of other integrable PDE is one of the most impressive achievements of the field of PDE in the last century.

⁴⁹I use this term here quite freely, it is typically used in a somewhat different context. Also some of the equations which I call phenomenological below, e.g dispersive equations, can be given formal asymptotics derivations

⁵⁰The system can be solved, in principle, by a well defined procedure which reduces the equation to a simple integration. A formal definition of integrability can be given for finite dimensional Hamiltonian systems and extended to some PDE, such as KdV.

It remains also the model case of a beneficial interaction between numerical experiments, heuristic applied mathematics arguments, algebra and rigorous analysis. Together they have led to the creation of a beautiful mathematical theory with extensive and deep applications outside the field of PDE where they have originated from. We have to be aware, however, of the obvious limitations of integrable systems; with few exceptions all known integrable evolution equations are restricted to one space dimension.

In the above discussion we have not mentioned diffusive equations⁵¹ such as the Navier-Stokes. They are in fact not variational and, therefore, do not quite fit in the above description. Though they could be viewed as phenomenological equations, they can also be derived from basic microscopic laws, such as those governing the mechanical, newtonian, interactions of a very large number of particles N . In principle⁵², the equations of continuum mechanics, such as Navier-Sokes, could be derived by letting the number of particles $N \rightarrow \infty$. Passing from discrete to continuous involves some loss of information hence the continuum equations have diffusive features. The best known examples of diffusive effects are the *heat conduction*, which appears in connection with the dissipation of energy in compressible fluids, and *viscosity*, corresponding to dissipation of momentum, in Fluids. Another example is that of *electrical resistivity* for the electrodynamics of continuum media. Particularly important in establishing a link between the microscopic, discrete, world of Newtonian particles and the continuous macroscopic ones described by Continuum Mechanics, is the Boltzmann equation.

As we have mentioned in section 1 diffusive equations appear as model equations in a wide range of applications to Engineering, Biology, Economics, Finance etc. They also appear naturally in connection to stochastic equations where, because of the complexity of the particular system involved, one needs to assume that various factors affecting the system are random.

Diffusive equations turn out to be also very useful in connection to geometric problems. Geometric flows such as mean curvature, inverse mean curvature, Harmonic Maps, Gauss Curvature and Ricci flows are some of the best known examples. Often these equations can be interpreted as the gradient flow for an associated elliptic variational problem.⁵³ They can be used to construct nontrivial stationary solutions to the corresponding stationary systems, in the limit as $t \rightarrow \infty$, or to produce foliations with remarkable properties, such as that used recently in the proof of the Penrose conjecture. As we have already mentioned, recently this idea has found an extraordinary application in the work of G. Perelman, who has used the Ricci flow to settle the three dimensional Poincaré conjecture.

⁵¹i.e. some of the basic physical quantities, such as energy, are not conserved and may in fact decrease in time. These are typically of parabolic type.

⁵²To establish this rigorously remains a major challenge.

⁵³One of the main new idea in the recent work of G. Perelman is to interpret Ricci flow as a gradient flow.

5. REGULARITY OR BREAK-DOWN

The most basic mathematical question in PDE is, by far, that of regularity of solutions. In the case of elliptic equations the issue is to determine the regularity of solutions to a geometric variational problem. In view of the modern way of treating elliptic equations, one first constructs a generalized solution by using the variational character of the equations⁵⁴. The original problem, then, translates to that of showing that the generalized solution has additional regularity. In many cases, such as the minimal hypersurfaces as graphs over mean convex domains in low dimensions, one can show that the generalized solutions are smooth. The same conclusion holds true, for example, for harmonic maps with target in a into a hyperbolic space. The solutions to the general Plateau problem⁵⁵, however, may have singularities. In this case the main issue becomes the structure of the singular sets of non-smooth solutions. Geometric measure theory provides sophisticated analytical tools to deal with this problem. Singularities are also known to occur in the case of higher dimensional harmonic maps, for positively curved target manifolds such as spheres.

In the case of evolution equations the issue is the possible spontaneous, finite time break-down of solutions, corresponding to perfectly nice initial conditions. This is a typical nonlinear, PDE phenomenon⁵⁶. It can be best illustrated in the case of the one dimensional Burger equation (39), $u_t + uu_x = 0$. As we have seen, all solutions, corresponding to smooth, compactly supported, nonzero initial data at $t = 0$, break-down in finite time. This despite the presence of infinitely many positive conserved quantities⁵⁷. Indeed one can show that the quantities $\int |u(t, x)|^{2k} dx$, $k \in \mathbb{N}$ are all conserved by the equation. The break-down corresponds, physically, to the formation of a shock wave. Similar examples of break-down can be constructed for compressible fluids or for some of the basic equations in elasticity. Singularities are also known to form, in some special cases, for solutions to the Einstein field equations in General Relativity. Moreover, one expects this to happen, in general, in the presence of strong gravitational fields. It is also widely expected that the general solutions of the incompressible Euler equations in three space dimensions, modeling the behavior of in-viscid fluids, break-down in finite time. Some speculate that the break-down may have something to do with the onset of turbulence for incompressible fluids with very high Reynolds numbers, described by the Navier-Stokes equations. In the case of Navier-Stokes the general consensus is that the evolution of all smooth, finite energy, initial data lead to global in time, smooth, solutions. This consensus has found an explicit formulation in one of the seven Clay Millennium Problems.

Break-down of solutions is also an essential issue concerning nonlinear geometric flows, such as the mean and inverse mean curvature flows, Ricci flow etc. As

⁵⁴The Euler Lagrange equations come in divergence form.

⁵⁵One looks at all minimal immersions, which may not be written as a graphs.

⁵⁶For smooth, one dimensional, Hamiltonian systems with positive energy, solutions are automatically global in time.

⁵⁷For ODE, even one positive conserved quantity would suffice to insure that solutions do not blow-up.

singularities do actually form in many important geometric situations, one is forced to understand the structure of singularities and find ways to continue the flow past them. Useful constructions of generalized flows can lead to the solution of outstanding geometric problems, as in the recent case of the Penrose conjecture in Riemannian geometry. The most important curvature flow in geometry, as we have already mentioned, turns out to be the Ricci flow. Recently G. Perelman has used the flow to prove the Poincaré conjecture in dimension three, one of the seven Millennium Clay problems. His work takes into account the fact that the Ricci flow may develop singularities, but avoids them by making *surgeries*, that is by removing regions of the compact three manifold near singularities. This has to be done in a controlled fashion and require the full power of a-priori estimates for the Ricci flow.

The problem of possible break-down of solutions to interesting, non-linear, geometric and physical systems is not only the most basic problem in PDE; it is also the most conspicuous unifying problem, in that it affects all PDE. It is intimately tied to the basic mathematical question of understanding what we actually mean by solutions and, from a physical point of view, to the issue of understanding the limits of validity of the corresponding physical theories. Thus, in the case of the Burger equation, for example, the problem of singularities can be tackled by extending our concept of solutions to accommodate *shock waves*, i.e. solutions discontinuous across curves in the t, x space. One can define, in this case, a functional space of generalized solutions in which the initial value problem has unique, global solutions. Though the situation for more realistic physical systems is far less clear and far from being satisfactorily solved, the generally held opinion is that shock wave type singularities can be accommodated without breaking the boundaries of the physical theory at hand. The situation of singularities in General Relativity is radically different. The type of singularities expected here is such that no continuation of solutions is possible without altering the physical theory itself. The prevailing opinion, in this respect, is that only a quantum field theory of Gravity could achieve this.

One can formulate a general philosophy to express our expectations with regard to regularity. To do that we need to classify our main equations according to the strength of their nonlinearities relative to that of the known *coercive* conservation laws or other a priori estimates. An estimate is called coercive if it allows one to derive size information about the solutions. For example in the case of the Laplace equation $\Delta u = f$ we can prove, by a straightforward integration by parts, the identity $\sum_{i,j} \int_{\mathbb{R}^n} |\partial_i \partial_j u|^2 = \int_{\mathbb{R}^n} |f|^2$, from which we can derive very useful size estimate for all second derivatives of u . Proceeding formally, exactly in the same manner, for the wave equation in \mathbb{R}^{1+1} $u_{tt} - u_{xx} = f$ we derive instead the useless identity $\int_{\mathbb{R}^{1+1}} (u_{tt}^2 - 2u_{tx}^2 + u_{xx}^2) = \int_{\mathbb{R}^{1+1}} |f|^2$ from which no size estimate for any of the second derivatives of u can be derived.

Among the basic conservation laws, discussed in connection to (119), that provided by the energy, defined by (132), is coercive, because it leads to an absolute, space-time bound on the size of solutions, or their first derivatives. The others, such as the linear and angular momentum, do not provide any additional informations

concerning local regularity. For most of the basic evolution equations, discussed in the previous section, the corresponding energy integral provides the best possible a priori estimate and therefore the classification is done relative to it.

In other cases, such as when there are additional symmetries, one often has better a priori estimates. For many elliptic and parabolic equations, for example, one can make use of the maximal principle or some monotonicity arguments to derive more powerful a priori estimates than those given by the energy integral. Integrable equations, such as KdV, also have additional, coercive, conservation laws. As explained above, the Burger equation has infinitely many positive conserved quantities. The incompressible Euler equations in dimension $n = 2$ have, in addition to the energy, a pointwise a priori estimate for the vorticity. It is for this reason that we can prove global regularity for 2D Euler equations. In all these cases the classification has to be done relative to the *optimal available a priori estimate*.

In what follows I will restrict myself to the case I find, personally, most interesting, that of the basic evolution equations for which there are no better, known, a priori estimates than those provided by the energy integral. These include all relativistic field theories, fluids, continuum mechanics and magneto-hydrodynamic, in three space dimensions and the absence of any additional symmetries. In these cases the classification is done by measuring the scaling properties of the energy integral relative to those of the equations. To illustrate how this is done consider again the nonlinear scalar equation $\square\phi - V'(\phi) = 0$ with $V(\phi) = \frac{1}{p+1}|\phi|^{p+1}$ mentioned earlier. Recall that the energy integral is given by (132). If we assign to the space-time variables the dimension of length, L^1 , then the spacetime derivatives have dimension L^{-1} and therefore \square has the dimension of L^{-2} . To be able to balance the left and right hand side of the equation $\square\phi = |\phi|^{p-1}\phi$ we need to assign a length scale to ϕ ; we find that to be $L^{\frac{2}{1-p}}$. Thus the energy integral, $E(t) = \int_{\mathbb{R}^d} (2^{-1}|\partial\phi|^2 + |\phi|^{p+1})dx$ has the dimension L^c , $c = d - 2 + \frac{4}{1-p}$, with d corresponding to the volume element $dx = dx^1 dx^2 \dots dx^d$ which scales like L^d . We say that the equation is *sub-critical* if $c < 0$, *critical* if $c = 0$ and *supercritical* for $c > 0$. Thus, for example $\square\phi - \phi^5 = 0$ is critical in dimension $d = 3$. The same analysis can be done for all our other basic equations. Yang Mills⁵⁸ is sub-critical for $n \leq 3$, critical for $n = 4$ and supercritical for $n > 4$. Wave maps⁵⁹ is sub-critical for $n = 1$, critical for $n = 2$, and supercritical for all other dimensions. The same holds true for the Einstein Vacuum equations. Most of our basic equations, such as Einstein-Vacuum, Euler, Navier- Stokes, Compressible Euler, Elasticity etc. turn out to be supercritical in the physical dimension $n = 3$.

A similar analysis can be done for the non-relativistic, nonlinear Schrödinger equation $i\partial_t\phi + \Delta\phi + (p+1)^{-1}|\phi|^{p-1}\phi = 0$, which has the conserved quantity $E(t) = \int_{\mathbb{R}^d} |\nabla\phi|^2 + |\phi|^{p+1}dx$. In this case, however, time and space scale differently. If we associate L^1 for each space variable, L^2 for the time variable and $L^{\frac{2}{1-p}}$ for ϕ , we find once more that E scales like L^c with $c = d - 2 + \frac{4}{1-p}$. Thus the critical exponents are the same as for the nonlinear wave equation.

⁵⁸In this case the gauge potential A scales like L^{-1} while F scales like L^{-2} .

⁵⁹The map ϕ scales like L^0 .

An evolutionary PDE is said to be regular if all smooth, finite energy, initial conditions lead to global smooth solutions. We expect that all sub-critical equations are regular while supercritical equations may develop singularities. Critical equations are important borderline cases. The heuristic reason is that the nonlinearity tends to produce singularities while the coercive estimates prevent it. In subcritical equations the coercive estimates is stronger while for supercritical equations it is the nonlinearity which is stronger. It may still be, however, that there are other more subtle a-priori estimates which are not taken into account by our crude heuristic argument. Thus, some supercritical equations, such as Navier-Stokes, may still be regular.

6. MAIN GOALS

The discussion of the previous sections suggests the following broad goals:

- 1.) *Understand the problem of evolution for the basic equations of Mathematical Physics.*
- 2.) *Understand in a rigorous mathematical fashion the range of validity of various approximations.*
- 3.) *Devise and analyze the right equation as a tool in the study of the specific geometric or physical problem at hand.*

We shall analyze them more closely in the following paragraphs.

6.1. Problem of evolution. The problem of evolution is embodied mathematically in the study of the initial value problem. The deterministic character of the fundamental equations of classical physics has its mathematical expression in the issue of existence and uniqueness of solutions to the corresponding Cauchy problem. One of the important achievements of the past century mathematics was the establishment of a general procedure which guaranties the existence and uniqueness of a local in time solution to broad classes of initial conditions and large classes of nonlinear equations, including all those we have already mentioned above. Yet this is just the beginning of the discussion. The long time behavior is far more difficult due, in particular, to the possible finite time break-down of solutions discussed in the previous section. Thus the most pressing issue at hand is to understand *when and how the local in time, smooth, solutions of our the basic equations develop singularities*. A simple-minded criterion for distinguishing between regular theories and those which may admit singular solutions is given by the classification between sub-critical and supercritical equations. As mentioned earlier, it is widely believed that *sub-critical equations are regular and that supercritical equations are not*. Indeed many sub-critical equations have been proved to be regular even though we lack a general procedure to establish it. The situation with supercritical equations

is far more subtle. To start with, an equation which we call now supercritical⁶⁰ may in fact turn out to be critical, or even sub-critical, upon the discovery of additional a-priori estimates. Thus an important question concerning the issue of criticality, and consequently that of singular behavior, is: *are there other, stronger, local a-priori bounds which cannot be derived from the Noether's principle ?* There are methods which can rule out the existence of some exact conserved quantities, different from the physical ones, yet there is no reason, I believe, to discount other, more subtle bounds. A well known *Morawetz multiplier* method leads, for some classes of nonlinear wave equations, to bounded space-time quantities which do not correspond to any conservation law. The Morawetz quantity, however, has the same scaling properties as the energy integral; it only provides additional information in the large. *The discovery of any new bound, stronger than that provided by the energy, for general solutions of any of our basic physical equations would have the significance of a major event.*

The critical equations, forming the boundary between the accessible sub-critical equations and the extremely difficult supercritical ones, have generated some of the most interesting mathematics in the last 20-30 years. Often they can be shown to be regular but this depends on the specific equation at hand. For example, it has been shown that the nonlinear wave equation $\square\phi - \phi^5 = 0$, which is critical in dimension $d = 3$, is regular for all initial data. The same problem for the critical nonlinear Schrödinger equation $i\partial_t\phi + \Delta\phi + \phi^5 = 0$ has turned out to be a lot more difficult and has only recently been cracked. There has also been a lot of progress on wave maps in the critical dimension $d = 2$. There is now hope that, in the case when the target manifold has constant negative curvature, i.e. hyperbolic space, wave maps in two space dimensions are also regular for all smooth data. the proof of these results require a remarkable interplay of geometric, i.e using the symmetries of the equation in physical space, and harmonic analysis methods, such as Strichartz and bilinear estimates and very refined micro-localization techniques. In elliptic theory there has been spectacular progress on critical nonlinear equations such, as the euclidean version of Yang -Mills equations in four dimensions, with deep applications⁶¹, due to Donaldson, to four dimensional topology.

Once we understand that the presence of singularities in our basic evolution equations is unavoidable we have to face the question of whether they can be somehow accommodated in a more general concept of solution, or whether their structure is such that the equation itself, indeed the physical theory which it underlies, becomes meaningless. An acceptable concept of generalized solution should, of course, preserve the deterministic nature of the equations in other words it should be uniquely determined from its Cauchy data.

Finally, once an acceptable concept of generalized solutions is found, we would like to use it to determine some important qualitative features, such as their large time

⁶⁰recall that what we call supercritical depends on the strongest a-priori, coercive, estimate available to us.

⁶¹Some of these applications have been recently obtained, using simpler equations than Yang-Mills, by Saiberg-Witten.

asymptotic behavior. One can formulate a limitless number of such questions, they are specific to each particular equation.

6.2. Range of validity of various approximations. The effective equations obtained by various limiting procedures or phenomenological assumptions can, of course, be studied in their own right just as those we have referred to above. They present us however with additional problems which have to do with the mechanics of their derivation from equations which we regard as more fundamental. It is entirely possible⁶², for example, that the dynamics of an effective system of equations leads to behavior *which is incompatible with the assumptions made in its derivation*. Or a particular simplifying ansatz, such as spherical symmetry in general relativity, or zero vorticity for compressible fluids, may turn out to be unstable in the large and therefore unreliable predictors of the general case. These and other similar situations lead to important dilemmas; should we persist to study the approximate equations nevertheless facing, in many cases, formidable mathematical difficulties⁶³, or should we abandon them in favor of the original system or a more suitable approximation? Whatever one may feel about this, in a specific situation, it is clear that the problem of understanding, rigorously, the range of validity of various approximations is one of the fundamental goals in PDE.

6.3. Devise and analyze the right equation. This last goal is equally important even though it is necessarily vague. The enormously important role played by partial differential equations in various branches of Mathematics is more evident than ever. One wonders in awe how equations such as Laplace, Heat, Wave Equation, Dirac, KdV, Maxwell, Yang-Mills, Einstein, which have been originally introduced in specific physical contexts, turned out to have such deep applications to seemingly unrelated problems in Geometry, Topology, Algebra or Combinatorics. Other partial differential equations appear naturally in Geometry when we look for embedded objects with optimal geometric shapes, such as isoperimetric problems, minimal surfaces, surfaces of least distortion or minimal curvature, or, more abstractly, connections, maps or metrics with distinguished properties. They are variational in character, just as the main equations of Mathematical Physics. Other equations have been introduced with the specific goal to allow one to deform a general object, such as a map, connection or metric to an optimal one. They come up, usually, in the form of geometric, parabolic flows. The most famous example of this is the Ricci flow, first introduced by R. Hamilton with the hope of using it as a way to deform Riemannian metrics to Einstein ones. Similar ideas were used earlier to construct, for example, stationary harmonic maps with the help of a harmonic heat flow or self dual Yang-Mills connections with the help of a Yang-Mills flow. As already mentioned, the Ricci flow has now been used successfully to settle the Poincaré conjecture in three space dimensions. Another remarkable, recent, example of the usefulness of geometric flows is that of the inverse mean flow, first introduced by Geroch, to settle the so called Riemannian version of the Penrose inequality.

⁶²See, for example, our discussion of infinite velocities for Navier-Stokes in the next section.

⁶³some which may turn out to be quite pathological, related, maybe, to the nature of the approximation.

7. GREAT OPEN PROBLEMS IN PDE

In this section I will state some of the main open problems in PDE in the form of five categories of conjectures. The first group of conjectures concern the Einstein vacuum equations, these are the famous Cosmic Censorship and Final State conjectures. The second group concerns the well known problem of break-down of smooth solutions for 3D Euler equations while the third deal with the problem of global regularity for 3D Navier-Stokes equations. The fourth group deals with the initial value problem for the supercritical nonlinear wave equation for which numerical experiments, suggest, just like for Navier-Stokes, that global regularity prevails for all regular data. Clearly the supercritical wave equation⁶⁴ is by no means equal in importance with the Einstein, Euler or Navier Stokes equations. I have included this problem among the others because, I believe, it is a great deal simpler than the regularity problem for Navier-Stokes. Its resolution, I believe, might open the way for understanding other supercritical equations. Finally the very loose Conjecture 5 concerns systems of conservation laws, such as Gas Dynamics, for which singularities are known to exist and thus the problem is to formulate a correct concept of generalized solution for which one can show existence and uniqueness. Despite my inability to give a more precise formulation, I have included it here in recognition of its enormous theoretical and practical importance.

Clearly other great problems exist which I do not have neither the space or expertise to discuss here. A glaring omission, for example, is that of the derivation of the the continuum equations of Hydrodynamics, such as those of Gas Dynamics or the Navier-Stokes equations, from the the basic principles of Newtonian mechanics. More precisely, Newtonian mechanics applies to a finite number of N interacting particles, corresponding to molecules, or atoms, in the atomistic conception of matter. One expects that, as $N \rightarrow \infty$, we can derive the continuum equations of motion of a gas or fluid. Here is the specific reference to this problem made by Hilbert in his famous address at the 1900 International Congress of Mathematicians,

Boltzmann's work on the principles of mechanics suggests the problem of developing mathematically the limiting processes, there merely indicated, which lead from the atomistic view to the laws of motion of continua.

There is no doubt however that the five categories of problems mention here are of fundamental importance.

7.1. Cosmic Censorship and Final State Conjectures. I start with a short introduction to the basic concepts needed to understand the statement of the conjectures, see also the compendium article concerning the Einstein equations. We shall consider here only the mathematical unambiguous case of Vacuum-Einstein manifolds⁶⁵. To solve the Einstein equations in vacuum we start with an initial data set which consists of a triplet $(\Sigma, g_{(0)}, k_{(0)})$ with Σ a three dimensional manifold,

⁶⁴One could have also included the supercritical nonlinear Schrödinger equation. The wave equation seems however easier to treat.

⁶⁵The conjectures, however, are suppose to hold for any physically reasonable matter-fields.

$g_{(0)}$ a Riemannian metric and $k_{(0)}$ a symmetric 2-tensor. In addition, one needs to impose a set of constraints on $g_{(0)}, k_{(0)}$ called *constraint equations*. To solve the Einstein vacuum equations amounts, roughly⁶⁶, to find a 3+1 dimensional manifold $\Sigma \subset M$ together with a Ricci flat (i.e. verifying (21)), Lorentz metric g on M such that $g_{(0)}$ is the restriction of g to Σ and $k_{(0)}$ is, roughly, the normal derivative⁶⁷ of g on Σ . The Lorentz manifold (M, g) is called an Einstein-vacuum (EV) development of the initial data set. Moreover we say that the development is a *Cauchy development* if every point of M , called events by physicists, is causally dependent of Σ , i.e. any physical observer⁶⁸ which experiences the event p , i.e. passes through p , must, at some point in the past or future of p , intersect Σ .

The simplest example of an EV manifold is the flat Minkowski space (\mathbb{R}^{1+3}, m) . We can view the hypersurface $t = 0$ as initial manifold Σ ; together with $g_{(0)} = e$ and $k_{(0)} = 0$, where e is the standard euclidean metric. In this way we get the flat initial data set $(\Sigma, e, 0)$, whose development is exactly the Minkowski space. We say that an initial data set $(\Sigma, g_{(0)}, k_{(0)})$ is asymptotically flat, if there exists a sufficiently large compact set of $K \subset \Sigma$ outside which the data set is close, in appropriate sense, to the flat one. The Minkowski metric can be written in spherical coordinates $t, r = |x|, \omega \in \mathbb{S}^2$ in the form, $-dt^2 + dr^2 + r^2 d\sigma_{\mathbb{S}^2}^2$ where $d\sigma^2$ represents the standard metric of the unit sphere \mathbb{S}^2 in \mathbb{R}^3 .

Another, very important, explicit solution of (EV) is given by the exterior Schwarzschild metric, of mass $m > 0$,

$$g_S = -\left(1 - \frac{2m}{r}\right)dt^2 + \left(1 - \frac{2m}{r}\right)^{-1}dr^2 + r^2 d\sigma_{\mathbb{S}^2}^2 \quad (136)$$

Though the metric seems singular at $r = 2m$ it turns out that one can glue together two regions $r > 2m$ and two regions $r < 2m$ of the Schwarzschild metric to obtain a metric which is smooth along $\mathcal{E} = \{r = 2m\}$, called the Schwarzschild horizon. The region $\{r > 2m\}$ is called the *domain of outer communication* of the Kerr space-time. It can be shown that the future and past sets of any point in this set intersects any timelike curve, passing through points of arbitrary large values of r , in finite time as measured relative to proper time along the curve. This fact is violated by points in the region $r \leq 2m$, which defines the *black hole* region of the space-time. Thus physical signals which initiate at points in $r \leq 2m$ cannot be registered by far away observers. Moreover the black hole region is singularity at $r = 0$.

The Schwarzschild metrics are a special case of a two parameter family of explicit metrics called Kerr. These metrics are stationary, which means, roughly, that the coefficients of the metric are independent of the time variable t .

The most primitive question asked about the initial value problem, which, as we have seen, has been solved for very large classes of evolution equations, is that of

⁶⁶The correct definition is done by an embedding of Σ to M

⁶⁷Derivative in the direction of the unit normal to Σ .

⁶⁸A physical observer is represented, mathematically, by time-like or null curve in M . This means a curve $\gamma(s)$ in M whose tangent vector $V(s) = \frac{d}{ds}\gamma(s)$ verifies $g(V(s), V(s)) \leq 0$ for every s along the curve.

local existence and uniqueness of solutions. For the Einstein equations this type of result was first established with the help of the wave coordinates mentioned in section 1, see (22), which allowed her to cast the Einstein equations in the form of a system of nonlinear wave equations to which one can apply the standard theory of nonlinear hyperbolic systems. The local existence result can actually be strengthened to a global result. This is done by proving that the local result can be extended, using the axiom of choice!, to a maximally extended Cauchy development.

Theorem *For each smooth initial data set there exists a unique maximal future Cauchy development.*

Thus any construction, obtained by an evolutionary approach from a specific initial data set, must be necessarily contained in its maximal development. This may be said to solve the problem of global⁶⁹ existence and uniqueness in General Relativity; all further questions may be said to concern the qualitative properties of the maximal Cauchy developments. The central issue becomes that of existence and character of singularities. First we can define a regular maximal development as one which is complete in the sense that all future time-like and null geodesics are complete. Roughly speaking this means that any freely moving observer in M can be extended indefinitely, as measured relative to its proper time. It turns out that any initial data set, which is sufficiently close to the flat one, admits a regular maximal Cauchy development.

Theorem[Global Stability of Minkowski] *Any asymptotically flat initial data set which is sufficiently close to the trivial one has a complete maximal future development. Moreover the curvature of the development is globally small and tends to zero at infinity, along any direction.*

At the opposite end of this result, when the initial data set is very far from flat, we have the following singularity theorem of Penrose,

Theorem *If the manifold support of an initial data set is noncompact and contains a closed trapped surface the corresponding maximal development is incomplete.*

The notion of a trapped surface $S \subset \Sigma$, can be rigorously defined in terms of a local condition on S . The flat initial data sets has, of course, no such surfaces. On the other hand, for the Schwarzschild initial data set, any surface $r = r_0$, with $r_0 < 2m$ is trapped. One can also check that the Schwarzschild metric has a genuine singularity at $r = 0$, where the curvature tensor becomes infinite. This is a lot stronger than just saying that space-time is incomplete. In fact the entire two parameter family of Kerr solutions, with the exception of the flat Minkowski space itself, have trapped surfaces and curvature singularities.

⁶⁹This is of course misleading, for equations defined in a fixed background global is a solution which exists for all time. In general relativity, however, we have no such background as the spacetime itself is the unknown. The connection with the classical meaning of a global solution requires a special discussion concerning the proper time of timelike geodesics.

The unavoidable presence of singularities, for sufficiently large initial data sets, has led Penrose to formulate two conjectures which go under the name of the *weak and strong cosmic censorship conjectures*. The first asserts that for all, but possibly an exceptional set of initial conditions, no singularities may be observed from infinity. Or, in other words, the singularities in General Relativity must be hidden, generically, by regions of space-time, called black-holes, in which all future causal geodesics remain necessarily trapped. To get a feeling for this consider the difference between the Minkowski space and the black hole region $\{r \leq 2m\}$ of a Schwarzschild space-time. In Minkowski space light originating at any point $p = (t_0, x_0)$ propagates, towards future, along the null rays of the null cone $t - t_0 = |x - x_0|$. Any free observer in \mathbb{R}^{1+3} , following a straight time-like curve, will necessarily meet this light cone in finite time, thus experiencing the physical event p . On the other hand, as we have mentioned earlier, any point p in the trapped region $r < 2m$ of the Schwarzschild space, is such that all the null rays initiating at p remain trapped in the black hole region $r < 2m$. The region $r > 2m$ is free of singularities, fortunately all singularities of Schwarzschild are such that they are invisible to the external region.

Given that singularities are, in general, unavoidable, it would be second to best⁷⁰ to have such a feature satisfied by all solutions of the Einstein equations. One would like to show, just like in a Schwarzschild space-time any *sufficiently distant observer* will never encounter singularities or any other effects propagating from singularities. To make this more precise one needs define what a sufficiently distant observer means. This is typically done by introducing the notion of future null infinity \mathcal{S}^+ which, roughly speaking, provides end points for the null geodesics which propagate to asymptotically large distances. The future null infinity is formally constructed by conformally embedding the physical spacetime M under consideration to a larger space-time \bar{M} with a null boundary \mathcal{S}^+ .

Definition. *The future null infinity \mathcal{S}^+ is said to be complete if any future null geodesics along it can be indefinitely extended relative an affine parameter.*

Given this enlarged space-time, with complete \mathcal{S}^+ , one defines the black hole region to be

$$\mathcal{B} = M - \mathcal{I}^-(\mathcal{S}^+) \tag{137}$$

with the chronological past $\mathcal{I}^-(\mathcal{S}^+)$ denoting all points that can be reached from \mathcal{S}^+ by past oriented time-like curves. The event horizon \mathcal{E} of the black hole is defined to be the boundary of \mathcal{B} in M . The requirement that space-time M has a complete future null infinity can be informally reformulated, by saying that the complement of the black hole region should be free of singularities. Indeed singularities outside the black hole region will necessarily affect the completeness of \mathcal{S}^+ . The black hole region, however, can only be defined a-posteriori after the completeness of \mathcal{S}^+ has been established.

⁷⁰The presence of singularities means that GR must fail. It would thus be comfortable to know that singularities are necessarily invisible to most observers and that outside black holes the theory holds true.

Here is now a more precise formulation of the Weak Cosmic Censorship (WCC) conjecture.

Conjecture 1.1 [*Weak Cosmic Censorship Conjecture (WCC)*] *Generic⁷¹ asymptotically flat initial data have maximal future developments possessing a complete future null infinity.*

Leaving aside the precise statement, to solve the conjecture amounts to find the boundary of the trapped region, called *event horizon*, and show that outside it the space-time is regular and behaves in fact like a perturbation of Minkowski space.

The WCC conjecture does not preclude the possibility that singularities may be visible by local observers, i.e. observers which may have already fallen inside a black hole. This could lead to causality violations, i.e. lack of unique predictability of outcomes of observations made by such observers. The strong cosmic censorship was designed to forbid such undesirable features of local singularities. According to it the maximal future Cauchy developments of an initial data set may not admit any extensions whatsoever, not only Cauchy developments. This would happen if the Lorentzian manifold itself cannot be extended, not only as solution of the Einstein equations, but as a regular Lorentz manifold. For example if we take away a point of \mathbb{R}^3 we obtain an incomplete but extendible manifold. A manifold with a cusp singularity⁷², however, cannot be extended beyond the cusp. Genericity is again important in view of some simple counterexamples, believed to be unstable.

Conjecture 1.2 [*Strong Cosmic Censorship (SCC)*] *Generic asymptotically flat or compact initial data sets have maximal future Cauchy developments which are locally in-extendible.*

The formulation above leads open the sense in which the maximal future developments are *in-extendible*. The precise notion of extendibility, which is to be avoided by SCC, is a subtle issue which can only be settled together with a complete solution of the conjecture. There have been various proposals among which I will only mention two.

- (1) The maximal future development is in-extendible as a $C^{1,1}$ Lorentzian manifold. This means, in particular, that some components of the curvature tensor must become infinite.
- (2) The maximal future development is in-extendible as a continuous Lorentzian manifold.

Though general, asymptotically flat, solutions of the Einstein vacuum equations are exceedingly complicated we expect that their asymptotic behavior is quite simple and is dictated in fact by the two parameter family of explicit Kerr solutions,

⁷¹There exist some pathological examples which exhibit *naked singularities*, i.e singularities which are not hidden by black holes and would therefore lead to a violation of the conjecture. For this reason we expect that weak cosmic censorship holds a set of initial conditions which is open and dense, in some sense, with respect to the set of all initial data.

⁷²with the curvature blowing up at the cusp.

corresponding to axially symmetric, rotating black holes. Here is a rough version of the conjecture.

Conjecture 1.3 [Final State Conjecture] *Generic asymptotically flat initial data sets have maximal future developments which can be described, asymptotically, as a finite number of black holes, Kerr solutions, moving away from each other.*

A simpler special case of this conjecture which would have to be understood long before the full conjecture is settled concerns the stability of just one Kerr solution:

Conjecture 1.3a [Global stability of Kerr] *Any small perturbation of the initial data set of a Kerr space-time has a global future development with a complete future null infinity which, within its domain of outer communication⁷³, behaves asymptotically like a (another) Kerr solution.*

At the present time we don't even have a satisfactory resolution of the issue of uniqueness of the Kerr solutions among all stationary solutions, see [H-E]. The only global stability result, among asymptotically flat space-times, we have today is that of the Minkowski space. The Cosmic Censorship conjectures have only been understood in special cases of solutions with a lot of additional symmetries.

7.2. Breakdown for 3D Euler. The incompressible Euler equations in \mathbb{R}^{d+1} take the form,

$$\partial_t u + u \cdot \nabla u = -\nabla p + f$$

where $f = (f_1, f_2, \dots, f_d)$ is a prescribed force, p is the hydrostatic pressure, $\nabla p = (\partial_1 p, \dots, \partial_d p)$ its gradient, and $u = (u_1, u_2, \dots, u_d)$ the velocity vector, satisfying the incompressible constraint

$$\operatorname{div} u = \sum_{i=1}^n \partial_i u_i = 0.$$

We are supposed to study the evolution of the initial data, $u(0, x) = u_0(x)$. For simplicity we shall take the case $f = 0$ and regular initial data $u_0 \in \mathcal{I}$; these are smooth data defined in the whole space which are either periodic or decay sufficiently fast at infinity.⁷⁴

Here are some of the few things we know about the 3D(d=3) Euler equations:

Theorem 2 (Local in time existence and uniqueness). *Given any regular initial data in $u_0 \in \mathcal{I}$ there exists a $T > 0$ and a unique solution⁷⁵ $u \in C^1([0, T] \times \mathbb{R}^3)$ of the incompressible Euler equations such that $u(0, x) = u_0(x)$.*

Theorem 3 (Unique Continuation). *Denote by $\omega = \operatorname{curl} u$ the vorticity of the fluid described by u . The local solution $u(t, x)$ of Theorem 1 can be uniquely continued past T as long as the integral $\int_0^T \|\omega(t)\|_{L^\infty} dt$ is finite.*

⁷³That means, roughly, its non-trapped region.

⁷⁴The initial data should have at least finite total energy and angular momenta. It is easy in fact to construct blowing-up solution with infinite energy.

⁷⁵One can show in fact that $u \in C^\infty([0, T] \times \mathbb{R}^3)$.

In particular this theorem, due to Beals-Kato-Majda, implies that the first possible singularity of the Euler flow must be accompanied by infinite vorticity. If we could prove that the vorticity remains bounded we would be able to extend the local existence result of Theorem 2.1 to infinite times. This is actually the case of 2D flows when the vorticity is conserved.

Conjecture 2.1 [*Breakdown of 3D Euler*] (weak form) *There exists a regular data $u_0 \in \mathcal{I}$, a time $T^* = T^*(u_0) > 0$ and a unique solution $u \in C^2([0, T^*) \times \mathbb{R}^3)$, $u(0, x) = u_0(x)$, such that $\|\omega(t)\|_{L^\infty} \rightarrow \infty$ as $t \rightarrow T^*$.*

Conjecture 2.2 [*Breakdown of 3D Euler*] (strong form) *Most regular initial data lead to such behavior. More precisely the subset of regular initial data $\mathcal{S} \subset \mathcal{I}$ which lead to finite time singularities must be dense with respect to the uniform norm L^∞ .*

An even stronger version of the conjecture ought to identify the complement of \mathcal{S} , i.e. $\mathcal{R} = \mathcal{I} \setminus \mathcal{S}$, as an exceptional set i.e. either having finite codimension or having, in a sense to be defined, measure zero. Given the exceptionally unstable character of the Euler equations it is in fact tempting to conjecture that the breakdown phenomenon is itself unstable which amounts to the fact that \mathcal{R} is itself dense in the set of all regular data \mathcal{I} .

Conjecture 2.3 [*Instability of Breakdown*] *The subset \mathcal{R} of regular initial conditions which lead to global in time C^2 solutions is dense in \mathcal{I} .*

7.3. Global Regularity for 3D Navier-Stokes. The Navier-Stokes equations in \mathbb{R}^{d+1} have the form,

$$\partial_t u + u \cdot \nabla u = \nu \Delta u - \nabla p + f$$

together with the incompressible constraint,

$$\operatorname{div} u = 0.$$

and the initial condition $u(0, x) = u_0(x)$. The notation for f, p, u are as for the Euler equations; $\nu > 0$ represents the viscosity. For simplicity we take once more the case $f = 0$ and regular initial data \mathcal{I} , defined just as for the Euler equations.

Here are the theorems analogous to Theorems 2 and 2

Theorem 4 (Local in time existence and uniqueness). *Given any smooth initial data $u_0 \in \mathcal{I}$ there exists a $T > 0$ and a unique solution⁷⁶ $u \in C^2([0, T] \times \mathbb{R}^3)$ of Navier-Stokes such that $u(0, x) = u_0(x)$.*

Theorem 5 (Unique Continuation). *The local solution $u(t, x)$ of Theorem 1 can be uniquely continued past T as long as the integral $\int_0^T \|u(t)\|_{L^\infty}^2 dt$ is finite.*

⁷⁶One can show in fact that $u \in C^\infty([0, T] \times \mathbb{R}^3)$.

In particular the local solutions to the Navier-Stokes equations can be continued as long as the velocity u remains bounded. Since large velocities do not seem to be generated either experimentally or numerically it is generally believed that the solutions to the 3D Navier-Stokes equations remain regular for all time. The belief is based upon numerical computations and theoretical inability to find any possible mechanism for blow-up. The following is one of the seven Clay millennium problems,

Conjecture 3.1[Global regularity for NS] *The solutions to the 3D Navier-Stokes equations, corresponding to any regular initial data $u_0 \in \mathcal{I}$, can be smoothly continued for all time.*

It is however entirely possible that singular solutions exist but are unstable and therefore difficult to construct analytically and *impossible to detect numerically or experimentally*. In that case a more realistic conjecture would be the following:

Conjecture 3.2[Generic Global regularity for NS] *There exists an exceptional⁷⁷ set $\mathcal{E} \subset \mathcal{I}$ of regular initial conditions, such all solutions with initial data $u_0 \in \mathcal{I} \setminus \mathcal{E}$ can be uniquely, smoothly, continued for all $t \geq 0$.*

Remark 8. Here is a short justification for Conjecture 3.2. Due to the supercritical character of the Navier-Stokes equations, and in the absence of any new stronger a priori estimates than that provided by the total energy, it is quite conceivable that singular solutions, corresponding to regular initial⁷⁸ data \mathcal{I} , do in fact exist. In view of Theorem 5 solutions of the Navier-Stokes equations can only break down if the velocity field becomes infinite. This situation is unphysical for many points of view.

- (1) Infinite velocities are physically unacceptable in view of the theory of relativity
- (2) An even more serious problem with large velocities is that of the incompressibility assumption. Indeed the incompressible equations are obtained as an approximation of the more general compressible equations whenever the particle velocities are far smaller than the sound speed of the fluid. This seems incompatible with very large velocities.
- (3) The Navier Stokes equations is also derived from well established continuum limits under a molecular assumption of *small mean free path* which seems, once more, incompatible with very large velocities.

Thus, if the breakdown actually occurs, we have to either give up on the claim that the Navier-Stokes equations provide a good approximation of physical reality or, luckily, all singular solutions turn out to be unstable and therefore Conjecture 3.2. is the correct formulation of what really happens.

⁷⁷of measure zero relative to an appropriate measure on \mathcal{I}

⁷⁸It is easy to construct blow-up solutions for initial data with infinite energy.

The next conjecture concerns the Leray solution. Leray was the first to construct global weak solutions of the Navier-Stokes equations corresponding to finite energy initial conditions. The problem with this type of solutions, called weak to distinguish them from classical and generalized solutions, is that they are derived in a non-dynamic fashion without concern for uniqueness, regularity or even energy conservation⁷⁹. For many years the hope has been that, by proving additional regularity properties, one can establish uniqueness of the Leray solutions and thus make them acceptable. After 70 years of fruitless attempts many of us doubt that the Leray solutions are reasonable solutions. It is quite conceivable in fact that all weak solutions for supercritical equations, derived non-dynamically with the help of the existing conservation laws(typically the energy) are not only non-unique but have also weird properties that make them completely unacceptable. An example of this type was given by V. Schaeffer; he has constructed weak solutions of the Euler equations in 2D which are compactly supported in space-time.

Conjecture 3.3[Weird properties of the Leray solution] *Show that the Leray solution is not unique, among all weak solutions with finite energy. Moreover show that there exist Leray solutions which exhibit weird, non-dynamical properties.*

8.1. Global Regularity for other supercritical equations. There exist other classes of supercritical equations, such as defocusing nonlinear Schrodinger and wave equations, for which one expects, based on numerical experiments, global regularity for all regular initial conditions. In what follows I will restrict my discussion to the case of scalar nonlinear wave equations of the form,

$$\square\phi - V'(\phi) = 0.$$

Here $\square = -\partial_t^2 + \partial_1^2 + \dots + \partial_d^2$ denotes the D'Alembertian in the Minkowski space \mathbb{R}^{d+1} and $V(\phi) = |\phi|^{p+1}$. We consider the initial value problem in \mathbb{R}^{d+1} ,

$$\phi(0, x) = \phi_0(x), \quad \partial_t\phi(0, x) = \phi_1(x).$$

We say that the initial data set ϕ_0, ϕ_1 is regular if both ϕ_0, ϕ_1 are smooth⁸⁰. As discussed in section 2 the equation is supercritical for $d > 2 + \frac{4}{p-1}$; if $d = 3$ this means $p > 5$. Here are the known facts:

Theorem 6 (Local in time existence and uniqueness). *Given any regular initial data there exists a $T > 0$ and a unique solution⁸¹ $\phi \in C^2([0, T] \times \mathbb{R}^3)$ of the nonlinear wave equation such that $\phi(0, x) = \phi_0(x)$, $\partial_t\phi(0, x) = \phi_1$.*

Theorem 7 (Unique Continuation). *The local solution $\phi(t, x)$ of Theorem 4.1 can be uniquely continued past T as long as the integral $\int_0^T \|\phi(t)\|_{L^\infty}^{p-3} dt$ is finite.*

The issue of global regularity for the sub-critical and critical cases is completely settled by the following:

⁷⁹The Leray solutions, like many others concepts of weak solutions for various nonlinear equations, satisfy only an energy inequality .

⁸⁰because of the finite propagation speed property of the wave equation we don't need to make assumptions about behavior at infinity.

⁸¹One can show in fact that $\phi \in C^\infty([0, T] \times \mathbb{R}^3)$.

Theorem 8 (Subcritical and critical global regularity). *If $p \leq 5$ the local solution of Theorem 4.1 can be continued for all $t \geq 0$.*

The supercritical case is entirely open. It is widely believed, based primarily on numerical experiments, that the following conjecture is true:

Conjecture 4.1[Supercritical global regularity] *The result of Theorem 4.3 can be extended to $p > 5$.*

To prove such a result, for all data, one needs a new a-priori estimate which would break the super-criticality of the problem. It is however conceivable that there are in fact plenty of unstable solutions, *impossible to detect numerically*, which break down in finite time; thus an alternative conjecture is the following:

Conjecture 4.2[Generic Global Regularity] *There exists an exceptional set of regular initial conditions whose evolution leads to finite time breakdown. All solutions, corresponding to data in the complement of the exceptional set, can be smoothly continued for all time.*

Remark 9. The possible existence of unstable initial conditions which lead to singularities makes the study of the supercritical wave equation relevant to the problem of global regularity for the Navier Stokes equations. The great advantage of the supercritical wave equation, by comparison to Navier Stokes, is that it is a scalar equation. Moreover it allows a very significant simplification to the case of solutions with spherical symmetry, in which case the problem becomes essentially one dimensional and the possible location of singularities is seriously constrained. Thus, unlike Navier-Stokes equations for which the problem of singularities has to be studied in full generality, the supercritical wave equations allows significant reductions which can be studied both theoretically and numerically.

9.1. Global, singular, solutions for 3-D Systems of Conservation Laws.

Many important, non-relativistic, equations of Continuum Mechanics, such as the Compressible Euler Equations, can be written as systems of hyperbolic conservation laws of the form:

$$\partial_t F^0(u) + \sum_{i=1}^3 \partial_i F^i(u) = 0.$$

Here $u = (u_1, u_2, \dots, u_N)$ and F_0, F_1, F_2, F_3 are smooth functions defined from some domain $\Omega \subset \mathbb{R}^N$ with values in \mathbb{R}^N , well specified in each particular case. We are supposed to solve the initial value problem $u(0, x) = u_0(x)$ with u_0 regular initial data. Here are the most basic facts known about such systems.

Despite the enormous applied interest of the basic conservation laws the mathematical theory remains highly unsatisfactory. There are however a few very important achievements which we outline below. We also refer to [Ma] for a thorough discussion.

- We have a well developed general theory of local in time existence and uniqueness of solutions to the basic conservation laws, based on the theory of symmetric hyperbolic systems.
- It is also well known that the solutions of all our basic systems of conservation laws may develop singularities in finite time. Such results are well understood in one space dimension where ∂u becomes infinite because of focusing of characteristics. Burger equation provides a very simple illustration of this phenomenon which is known as formation of shock waves. Less satisfactory results are also known in higher dimensions.
- A satisfactory notion of global generalized solution, based on the bounded variation norm, exists in 1D. This allows us to prove a global existence and uniqueness result for general initial data with small total variation for a large class of 1D systems of conservation laws.

It is impossible to give in a short presentation a precise mathematical statement for the Conjecture below.

Conjecture 5.1[*Existence and Uniqueness of Generalized Solutions*] *In the case of specific physical systems, such as compressible gas dynamics and nonlinear elasticity, one can define an appropriate concept of generalized solution, compatible with shock waves and other possible singularities, for which we can prove global existence and uniqueness of the initial value problem corresponding to all, or generic, regular initial conditions.*

Remark. To illustrate the enormous difficulty of this conjecture we remark that full treatment of the Compressible Euler equations in 3D, for example, must include the limiting case of the incompressible equations. This requires not only to settle first the break-down Conjecture 2.1. but also to find a way of continuing the solutions of the incompressible Euler equations past singularities. Obviously there is no hope at this time that this could be achieved soon. This illustrates the need to work on vastly simplified model problems.

9.2. Conclusions: Here is a short list of remarks concerning the problems discussed above.

I. *All seem inaccessible at the present time.*

II. *Though each problem is different and would ultimately require the development of custom-tailored techniques they share important common characteristics:*

- (1) They are all supercritical
- (2) They all seem to require the development of generic methods which allow the presence of exceptional sets of data. The development of such methods may be viewed as one of the great challenges for our century.
- (3) Conjectures 1,4,5 require the development of a powerful hyperbolic theory comparable with the progress made last century in elliptic theory.

III. *Need to concentrate on simplified model problems.*

There are plenty of great simplified model problems in connection with Cosmic Censorship, based on assuming special symmetries which significantly reduce the complexity of the Einstein equations. Conjectures 4 and 5 allow also for important simplifications. Conjectures 2 and 3, however, seem irreducible hard, by which I mean that it is difficult to find simpler model problems whose resolution would significantly help us understand the original ones.

REFERENCES

- [And] L. Andersson *The global existence problem in General Relativity*, preprint.
- [Br-Bro] *Partial Differential equations in the 20th century* Advances in Math. **135** (1998), p. 76-144.
- [Br] Y. Choquet-Bruhat *Theoreme d'Existence pour certains systemes d'equations aux derivées partielles nonlineaires.*, Acta Math. **88** (1952), 141-225.
- [Chr] D. Christodoulou *On the global initial value problem and the issue of singularities*, Class. Quant. Gr.(1999) A23-A35.
- [H-E] S. Hawking, G.F.R. Ellis *The large scale structure of space-time*, Cambridge University Press, Cambridge 1973.
- [Ch-Kl] D. Christodoulou and S. Klainerman, *The global nonlinear stability of the Minkowski space*, Princeton University Press (1993) 1986
- [John] F. John *Partial Differential Equations* fourth edition, Springer.
- [Kl1] S. Klainerman *PDE as a unified subject* Special volume GAFA2000.
- [Kl2] S. Klainerman *Mathematical Challenges of General Relativity* The ninth Marcel Grossman Meeting ed. V.G Gurzadyan, R.T. Jantzen , R. Ruffini, pages 28-41, 2000.
- [Pe] R. Penrose, *Singularities and time asymmetry in General Relativity-an Einstein centenary survey*, S. Hawking, W. Israel, , Cambridge Univ. Press, Cambridge 1979.
- [Ma] A. Majda, *Compressible Fluid Flow and System of Conservation Laws in several space variables*. Applied Math. Sciences 53, Springer-Verlag 1984.
- [Wa] R. Wald *General Relativity* University of Chicag Press, 1984

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON NJ 08544

E-mail address: `seri@math.princeton.edu`