GENERALIZED SOLUTIONS

TERENCE TAO

1. Generalized solutions

In many applications of mathematics, one uses a set of equations (often a set of partial differential equations) to model some system in real-life or in theoretical science; a typical such system might be expressible as a *linear partial differential equation*

$$Lu(x) = f(x) \text{ for all } x \in \Omega$$

where Ω is the domain of space (or spacetime), u represents the unknowns (either scalar or vector-valued), L is some linear differential operator, and f is some given data, though certainly more complicated nonlinear equations are certainly possible[1]. However, in some cases the assumptions used to justify the model are not uniformly valid, but may break down (or become otherwise dubious) at isolated places. For instance, an assumption that a certain field or medium is continuous may fail if there is a *phase transition* (e.g. a transition between liquid and solid matter) or a *singularity* (e.g. point masses or point charges with infinite mass density or charge density). In such cases, there will be locations where the various functions used to describe the system are not differentiable, and hence the partial differential equation $Lu = f$ does not make sense. Nevertheless, one would still like to have a mathematically meaningful notion of a *solution* to a partial differential equation, even in the presence of singularities. This leads to various notions of a *generalized solution*; a major task in this field is to identify to what extent these notions are compatible, a task which is now quite well understood for linear equations such as $Lu = f$ but still very interesting for non-linear equations.

One possible notion of a generalized solution to an equation such as $Lu = f$ is to allow for the existence of some *singular set* $S \subset \Omega$ in which the solution u is allowed to be singular or undefined, but require that u be smooth outside of S (or at least smooth enough that it is clear how to define Lu), and only require that the equation $Lu(x) = f(x)$ be true outside of S. Typically the set S will be closed and suitably "small" (e.g. zero measure, or having positive codimension, or being contained in a finite union of hypersurfaces). This is one of the simplest ways to define a generalized solution, but it is usually too weak in the sense that it does not exclude various physically implausible solutions. For instance, consider an equation describing a field (e.g. an electromagnetic field) permeating two types of material (say glass and air), with a discontinuity at the boundary S between

[1]For simplicity we shall completely ignore the important issue of what kind of *boundary conditions* to specify on the boundary of Ω.

the glass and the air. One might have one equation for the field inside the glass, and one equation for the field inside the air. But if we do not require any sort of "compatibility condition" at the boundary S then there is no relationship between the two different materials, and we will not be able to answer such basic questions as what happens to an electromagnetic wave in air when it encounters glass. This is clearly unsatisfactory from the viewpoint of applications, and so we seek better notions of generalized solution[2] which have some non-trivial content at singularities.

There are a number of ways to achieve this, often based from motivation from the physical world. For instance, in the real world, we expect there to be no actual singularities (except perhaps in extreme situations such as black holes), only approximate singularities, where for instance a function gets very large or very steep but not actually infinite or non-differentiable. One can incorporate this in the mathematical model by introducing some sort of *regularizing parameter* $\varepsilon > 0$ which has the effect of automatically truncating all singularities past a certain cutoff, constructing solutions for the regularized model, and then taking limits as $\varepsilon \to 0$. (Variants of this method include viscosity methods, penalization methods, and finite difference schemes). For instance, in order to avoid the singularity in Newton's famous inverse square law $F_{x,y} = -\frac{Gm_1m_2(x-y)}{|x-y|^3}$ for gravitation at the diagonal $x = y$, one might regularize this law by instead considering a law of the form $F_{x,y} = -\frac{Gm_1m_2(x-y)}{(\varepsilon^2+|x-y|^2)^{3/2}}$, which avoids singularity at the origin but approaches the original law when $\varepsilon \to 0$. Mathematically, this might result in a family u_ε of smooth solutions to an approximate equation, such as $L_\varepsilon u_\varepsilon = f_\varepsilon$, such that f_ε converges in some suitable sense to f and (the coefficients of) L_ε converge to (the coefficents of) L. The mathematical task is then to extract some limit u out of u_ε and establish what properties it has; typically it will solve the original equation $Lu = f$ not in the classical sense (because u might not be smooth), but in some weaker sense, for instance in the sense of distributions or away from a singular set. In some situations, these notions of solutions enjoy uniqueness and existence (for the given data f, there is a unique u which solves the equation in this sense), which is a good test to indicate that this is a workable notion of solution. A good example is the *inviscid Burgers equation* $\partial_t u + u\partial_x u = 0$, for which uniqueness of solutions breaks down once discontinuities appear; however, by working first with the *viscous Burgers equation* $\partial_t u - \varepsilon\partial_{xx} u + u\partial_x u = 0$ and then taking limits as $\varepsilon \to 0$ we retain uniqueness of solutions even after the first singularity develops. However, in other situations, the convergence of u_ε to u is only in some very weak sense (e.g. in the sense of distributions), and this can cause undesirable and non-physical consequences (e.g. the mass or energy of u may disappear into a singularity, and then violate a physical conservation law). Sometimes one has to add such conservation laws as an additional hypothesis in order to obtain uniqueness and existence; similarly for the other notions of generalized solution we discuss below.

A related way to obtain a good concept of generalized solution is not to view the partial differential equation $Lu = f$ as the fundamental object generating the

[2]There are certain classes of equations, notably *elliptic* equations, in which singularities are automatically guaranteed to not be present; a famous example is Laplace's equation $\Delta u = 0$, whose solutions are harmonic and hence smooth. In such cases the notions of weak solution and strong solution usually co-incide.

model, but rely instead on some other characterization. One common characterization is a *variational* characterization; for instance, a solution u might be defined as a minimizer (or at least a critical point) of a *Lagrangian functional* such as

$$\mathcal{L}(u) := \int_\Omega F(u, \nabla u, \ldots, \nabla^k u, f)$$

for some explicit function F of the data f, and of u and its first few derivatives; often one also constrains u and its derivatives to obey various physical conditions (e.g. a field which is supposed to represent mass density might be constrained to be non-negative). When u and f are smooth, calculus of variations allows one to rewrite this variational problem as an *Euler-Lagrange equation*, which is usually a partial differential equation. However, the rigorous justification of this equation is not automatic if u or f has singularities. Instead, what one can do is add a *penalty term* to the Lagrangian $\mathcal{L}(u)$, creating a new Lagrangian $\mathcal{L}_\varepsilon(u)$. Typically, this penalty term is designed to "penalize" any severe oscillation or singularity of u by making $\mathcal{L}_\varepsilon(u)$ extremely large when u exceeds certain thresholds, for instance if the size of u exceeds $1/\varepsilon$. This usually is enough to ensure that the minimizers u_ε of the Lagrangian exist, are unique, and are smooth. One then takes limits as $\varepsilon \to 0$; usually one designs the penalized Lagrangians such that $\mathcal{L}_\varepsilon$ converges (formally, at least) to $\mathcal{L}$, but often to pass from this to a statement that u_ε converges to some suitable u requires a non-trivial amount of analysis.

The above approaches to defining generalized solutions were based on approximating the rough putative solution by smoother objects which were already known to be solutions, or at least to be approximate solutions. Another approach is simply to accept that the solution u contains singularities, and make sense of the partial differential equation such as $Lu = f$ (or the variational characterization) in a way which does not require u to be smooth. For instance, one might only require Lu to equal f in the sense of distributions, which means that

$$\langle u, L^*\phi \rangle = \langle f, \phi \rangle \text{ for all test functions } \phi$$

where L^* is the adjoint operator to ϕ (so $\langle u, L^*\phi \rangle = \langle Lu, \phi \rangle$ for smooth functions u at least). Another possibility is to rewrite $Lu = f$ in a form which involves smoother operations, such as[3] $u = L^{-1}f$. For instance, the simplest differential equation,

$$\frac{d}{dx}u(x) = f(x),$$

ostensibly only makes sense if u is differentiable. However, if we rewrite this differential equation (using t he fundamental theorem of calculus) as an integral equation

$$u(x) = u(0) + \int_0^x f(y)\ dy$$

then this weaker notion of a solution to $\frac{d}{dx}u = f$ does not necessarily require u to be differentiable, instead merely requiring u to be continuous and f to be

[3]In many cases the inverse of L is not well understood (otherwise solving $Lu = f$ would not be so difficult in the first place!). Instead, one often approximates L by a simpler operator L_0 whose inverse *is* well understood, and rewrite $Lu = f$ as $L_0 u = (L_0 - L)u + f$ and hence $u = L_0^{-1}((L_0 - L)u + f)$. Clearly many further variations on this theme are possible; the idea of learning about the PDE theory of L through the PDE theory of an approximant L_0 is part of *perturbation theory*, which is an important component of analysis in PDE.

conditionally integrable, and thus allows for more general notions of solutions (e.g. $u(x) = |x|$ and $f(x) = \mathrm{sgn}(x)$). This trick of converting a differential equation into an integral equation (which usually requires much less regularity) is a very common and powerful way of defining a workable notion of generalized solution. Sometimes there is a difficulty in that the operator L^{-1} is not uniquely defined; for instance, if L is self-adjoint, then usually one can define inverses $(L + i\varepsilon)^{-1}$ and $(L - i\varepsilon)^{-1}$ for any $\varepsilon > 0$, but the two families of operators can converge to different limits as $\varepsilon \to 0$. Such subtleties often require a great deal of care and analysis.

Similarly, we can often use the Lagrangian formulation to define solutions even in the presence of singularities. For instance, if u has a singular set S, we can sometimes still define the Lagrangian simply by excluding S, thus

$$\mathcal{L}(u, S) := \int_{\Omega \backslash S} F(u, \nabla u, \dots, \nabla^k u, f).$$

Then there are two notions of a minimizer or critical point of $\mathcal{L}$; either one fixes S and only allows u to vary (and be smooth) outside of S, or one allows both u and S to vary (e.g. by applying a smooth deformation of the domain to move S). The distinction between the two concepts of solutions (sometimes denoted *critical points* and *stationary solutions* respectively) can be quite subtle, and is especially important in certain geometric PDE such as harmonic maps and Yang-Mills equations.

Department of Mathematics, UCLA, Los Angeles CA 90095-1555

E-mail address: tao@math.ucla.edu