Chapter 9
Relationships Between Different Partial Differential Equations

9.1 The Continuity Equation for a Dynamical System

As in Sect. 6.1, we consider a system of ordinary differential equations (ODEs)

$$\frac{\mathrm{d}x^i(t)}{\mathrm{d}t} = F^i(t, x^1(t), \ldots, x^d(t)), \text{ for } i = 1, \ldots, d. \tag{9.1.1}$$

For notational convenience, we shall leave out the vector index i; thus, in the sequel. x may stand for the vector $(x^1, \ldots, x^d)$.

We assume that F in (9.1.1) is Lipschitz continuous with respect to x and continuous with respect to t so that the Picard–Lindelöf theorem guarantees the existence of a solution for $0 \le t \le T$ for some $T > 0$. In the sequel, we shall often assume that this holds for $T = \infty$.

We let $x(t, y)$ be the solution of (9.1.1) with

$$x(0, y) = y. \tag{9.1.2}$$

The idea is to consider the flow generated by the system (9.1.1). That is, for a measurable set $A \subset \mathbb{R}^d$ of initial values, we consider the set $A_t := x(t, A)$ of their images under the dynamical system (9.1.1). Instead of sets, however, it is more useful to consider probability densities $h(x, t)$ of x; that is, for each measurable $A \subset \mathbb{R}^n$, the probability that $x(t)$ is contained in A is given by

$$\int_A h(y, t)\mathrm{d}y, \tag{9.1.3}$$

and we have the normalization

$$\int_{\mathbb{R}^d} h(y, t)\mathrm{d}y = 1 \text{ for all } t \ge 0. \tag{9.1.4}$$

J. Jost, *Partial Differential Equations*, Graduate Texts in Mathematics 214,
DOI 10.1007/978-1-4614-4809-9_9,

When x satisfies (9.1.1), the density h, as a quantity derived from x, then also has to satisfy some evolution equation. In fact, h evolves according to the continuity equation

$$\frac{\partial}{\partial t}h(x,t) = \sum_{i=1}^{d} \frac{\partial}{\partial x^i}(-F^i(t,x)h(x,t)) = -\mathrm{div}(hF). \tag{9.1.5}$$

This equation states that the change of the probability density in time is the negative of the change of the state as a function of its value. (In mechanics, this is also called the conservation of mass equation. It represents the Eulerian point of view that works with fields in contrast to the Lagrangian point of view that considers the individual trajectories of (9.1.1).) We note that this equation is a generalization of the Eq. (8.2.44) derived in Sect. 8.2.

We shall give two derivations of (9.1.5). At this point, these derivations will be formal, in the sense that we do not yet know whether a solution exists. Actually, the existence issue has already been addressed in Sect. 7.2, as we shall remark below.

We consider the functional

$$\begin{aligned} I(A,\epsilon) :&= \int_{A_{t+\epsilon}} h(x(t+\epsilon,y),t+\epsilon)\mathrm{d}x(t+\epsilon) \\ &= \int_{A_t} h(x(t+\epsilon,y),t+\epsilon) \det \frac{\partial x(t+\epsilon)}{\partial x(t)} \mathrm{d}x(t), \end{aligned}$$

where in the last step, we have used the flow x to transform the set $A_{t+\epsilon}$ back into the original set A_t. We compute (with some obvious shorthand notation)

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} I(A,\epsilon)_{|\epsilon=0} = \int_{A_t} (h_t + h_x x_t + h\ \mathrm{div} x_t)\mathrm{d}x = \int_{A_t} (h_t + \mathrm{div}(hF))\mathrm{d}x, \tag{9.1.6}$$

where, of course, for the last step, we have used (9.1.1). Since this holds for every A, (9.1.5) follows.

For the alternative derivation of (9.1.5), we assume that we have an initial density

$$\eta(x) := h(x,0), \tag{9.1.7}$$

and we write

$$h(x,t) =: Q_t \eta(x). \tag{9.1.8}$$

This indicates that we consider the density $h(t,.)$ as the temporal evolution of the initial density η under the dynamical system (9.1.1). Equation (9.1.3) then yields

$$\int_A Q_t \eta(x) = \int_{x(t,.)^{-1}(A)} \eta(x) = \int_{\mathbb{R}^n} \eta(x)\chi_A(x(t,x)) \tag{9.1.9}$$

for the characteristic function χ_A of the set A.

When we put, for a function ψ, $k(t, x) := \psi(x(t, x))$, we have

$$k(t, x(-t, x)) = \psi(x) \tag{9.1.10}$$

(because $x(t, x(-t, x)) = x$). Taking the total derivative of (9.1.10) with respect to t then yields at $t = 0$, and hence by time translation at every $t \geq 0$,

$$\frac{\partial k(t, x)}{\partial t} - \sum_i \frac{\partial k(t, x)}{\partial x^i} F^i(t, x) = 0. \tag{9.1.11}$$

Inserting (9.1.11) with $\psi = \chi_A$ into (9.1.9) yields

$$\begin{aligned}\int_A \frac{\partial}{\partial t} Q_t \eta(x) &= \int \eta(x) \sum_i \frac{\partial \chi_A(x(t, x))}{\partial x^i} F^i(t, x) \\ &= -\int \sum_i \frac{\partial}{\partial x^i}(\eta(x) F^i(t, x)) \chi_A(x(t, x)) \\ &= -\int_A Q_t \left(\sum_i \frac{\partial}{\partial x^i}(\eta(x) F^i(t, x)) \right).\end{aligned}$$

Since this holds for every measurable $A \subset \mathbb{R}^d$, we see, recalling (9.1.8), that $h(x, t)$ satisfies (9.1.5), indeed.

Equation (9.1.5) is, of course, the same as (7.2.5), the partial differential equation of first order that we have studied in Sect. 7.2.

9.2 Regularization by Elliptic Equations

As already emphasized repeatedly, a crucial issue in the theory of PDEs is regularity of solutions. We have to break the circulus vitiosus that in order to qualify as a solution of some PDE, a function should be sufficiently differentiable, but a PDE can force any putative solution to have some singularities, and the spaces in which we may naturally seek solutions and the schemes by which we try to obtain them typically also contain nonsmooth functions. We have already seen the basic idea how to overcome this problem, namely, to relax the requirement for a function to count as a solution. More precisely, we seek some criterion that, for a differentiable function, is necessary and sufficient to be a solution of the PDE in question, but that as such does not depend on the differentiability of that function. A function that then satisfies this requirement, without necessary being differentiable, is called a weak solution of that PDE. The existence problem for solutions of PDEs is thereby broken up into two subproblems. The first one concerns the existence of a weak solution, and the second one consists in the investigation of the regularity properties of weak solutions. For certain classes of PDEs, in particular, elliptic ones, as we shall see in subsequent chapters, one can show that any weak solution is sufficiently differentiable. In that case, the scheme then succeeds in finding a classical solution. In other cases, weak solutions may inevitably have some singularities. One then tries to understand the nature of these singularities and what constraints weak solutions have to obey.

There exist two important methods for defining weak solutions. One, which we shall explore in Sects. 10.1 and 11.2, simply multiplies the differential equation in question by any smooth functions, so-called test functions, and integrates by parts to shift the derivatives from the unknown, and perhaps singular, solution to the test functions. When the resulting identity is satisfied for all test functions, we have a weak solution. The second method is based on the observation that solutions of many PDEs have to satisfy some maximum principle and, conversely, can be characterized by that maximum principle. Again, the maximum principle by itself does not stipulate any differentiability, and therefore, one can try to develop a concept of weak solution on the basis of the maximum principle. In fact, as we shall explain, the maximum principle can even achieve more than that. It can yield a selection principle among possible weak solutions, or expressed differently, it can enforce uniqueness of weak solutions by selecting that weak solution that is best possible in the sense of regularity properties.

Let us describe the idea first before we implement it in an existence scheme.

We recall from Chap. 2 that a twice differentiable function g is called harmonic in the domain $\Omega \subset \mathbb{R}^d$ if

$$\Delta g(x) = 0 \text{ for all } x \in \Omega. \tag{9.2.1}$$

Likewise, such a function γ is called subharmonic in Ω if

$$\Delta \gamma(x) \geq 0 \text{ for all } x \in \Omega. \tag{9.2.2}$$

It turns out that these two concepts, harmonic and subharmonic, can be defined in terms of each other, so as to dispense with the smoothness requirements. The property required in the following definition is equivalent to (9.2.2) when γ is twice differentiable.

Definition 9.2.1. Let $\gamma : \Omega \to [-\infty, \infty)$ be upper semicontinuous (i.e., whenever $(x_n) \subset \Omega$ converges to $x \in \Omega$, then $\gamma(x) \geq \limsup_{n\to\infty} \gamma(x_n)$), with $\gamma \not\equiv -\infty$. Then γ is called subharmonic in Ω if whenever g is harmonic in $\Omega' \Subset \Omega$ and $\gamma \leq g$ on $\partial\Omega'$, then also

$$\gamma \leq g \text{ in } \Omega'. \tag{9.2.3}$$

Thus, a not necessarily smooth subharmonic function can be characterized in terms of harmonic functions. A subharmonic function has to lie below any harmonic function with the same boundary values on some subdomain of Ω. Conversely, we would like to characterize a harmonic function by always lying above subharmonic functions with the same boundary values. This is not yet sufficient, however, because this property also holds for any superharmonic function (superharmonic functions are defined in the same as subharmonic ones, simply by reversing inequalities; for instance, in the smooth case, a subharmonic γ has to satisfy $\Delta\gamma \geq 0$). Of course, we could then characterize a harmonic function by lying above all subharmonic and below all superharmonic ones. It is often more convenient, however, to use only subharmonic function and obtain a harmonic function as the smallest function that lies above all subharmonic ones. That is the idea of the Perron method that we have developed in Sect. 4.2. We recall Theorem 4.2.1.

Theorem 9.2.1. *Let ϕ be a bounded function on $\partial\Omega$. Then*

$$g(x) := \sup_{\gamma \leq \phi \text{ on } \partial\Omega, \gamma \text{ subharmonic in } \Omega} \gamma(x) \tag{9.2.4}$$

is a harmonic function on Ω. g is smooth in Ω. Under suitable regularity conditions on Ω and ϕ (not specified here), it satisfies the Dirichlet condition $g(y) = \phi(y)$ for $y \in \partial\Omega$.

This, however, in the present context is only an interlude, meant to motivate a solution concept for certain first-order equations where, according to our considerations above, we need to reckon with singularities as well as with issues of non-uniqueness. We consider problems of the form

$$\begin{aligned} \frac{\partial h(x,t)}{\partial t} + J\left(\frac{\partial h(x,t)}{\partial x}, x\right) &= 0 \text{ for } x \in \mathbb{R}^d, t > 0 \\ h(x,0) &= h_0(x) \text{ for } x \in \mathbb{R}^d. \end{aligned} \tag{9.2.5}$$

Here, J is assumed to be bounded and continuous. Equation (9.2.5) can be seen as a generalization of (7.2.14), and therefore, in particular, we have to reckon with all the phenomena discussed in Sect. 7.2. As argued there, in general, we cannot expect to find a differentiable solution of (9.2.5), and on the other hand, when we give up the smoothness requirement, there could be several functions that may count as a solution. Thus, we wish to find among those a best one. Of course, we need to qualify what “best” means here. For instance, it could select a solution that is most regular or, put differently, has the mildest possible singularity.

The idea of viscosity solutions that has been developed in [6] and that we are going to present now consists in approximating (9.2.5) by another equation with better solution properties. We then take the solutions of the approximating equations and hope that they tend to a good solution of the original equation (9.2.5) when the approximation parameter goes to 0. This works as follows:

$$\begin{aligned} \frac{\partial h^\epsilon(x,t)}{\partial t} + J\left(\frac{\partial h^\epsilon(x,t)}{\partial x}, x\right) &= \epsilon\Delta h^\epsilon \quad \text{for } x \in \mathbb{R}^d, t > 0 \\ h^\epsilon(x,0) &= h_0(x) \quad \text{for } x \in \mathbb{R}^d. \end{aligned} \tag{9.2.6}$$

$\epsilon > 0$ is the approximation parameter that we want to let to tend to 0. The term $\epsilon \Delta h^\epsilon$, while being of higher order than the terms in the original equation (9.2.5), ensures the regularity of solutions. (9.2.6) is a parabolic equation, and its solutions h^ϵ satisfy some maximum principle as we shall explain and utilize below. The idea then is to obtain or define a solution of (9.2.5) as $h = \lim_{\epsilon \to 0} h^\epsilon$ (perhaps, we might have to take a subsequence, but this is not important for the principal idea). h need not be smooth or even continuous, even though the h^ϵ are, because the regularity properties of the latter may become worse and worse as $\epsilon \to 0$. Nevertheless, certain properties of the h^ϵ should persist in the limit. In fact, it turns out that h can be characterized and distinguished from other solutions of (9.2.5) by some maximum principle property which we shall now explain. The key point is that this is a property of the approximation solutions that does not depend on the value of $\epsilon > 0$.

Suppose that

$$h^\epsilon - \eta \text{ has a maximum at } (x_0, t_0) \in \mathbb{R}^d \times (0, \infty) \tag{9.2.7}$$

for some smooth function η. Then

$$\frac{\partial h^\epsilon(x_0, t_0)}{\partial t} = \frac{\partial \eta(x_0, t_0)}{\partial t}, \quad \frac{\partial h^\epsilon(x_0, t_0)}{\partial x} = \frac{\partial \eta(x_0, t_0)}{\partial x} \tag{9.2.8}$$

and

$$\Delta h^\epsilon(x_0, t_0) \leq \Delta \eta(x_0, t_0). \tag{9.2.9}$$

Therefore, from (9.2.6), we can deduce that

$$\frac{\partial \eta(x_0, t_0)}{\partial t} + J\left(\frac{\partial \eta(x_0, t_0)}{\partial x}, x\right) \leq \epsilon \Delta \eta(x_0, t_0). \tag{9.2.10}$$

A key point here is, of course, that Eq. (9.2.5) involves only first derivatives of h, but not h itself.

Similarly, when $h^\epsilon - \tilde{\eta}$ has a minimum at (x_0, t_0), we have

$$\frac{\partial \tilde{\eta}(x_0, t_0)}{\partial t} + J\left(\frac{\partial \tilde{\eta}(x_0, t_0)}{\partial x}, x\right) \geq \epsilon \Delta \tilde{\eta}(x_0, t_0). \tag{9.2.11}$$

Conversely, when these inequalities hold for any such η or $\tilde{\eta}$, resp., then h^ϵ is a solution of (9.2.6). The expectation that this property passes to the limit $\epsilon \to 0$ now motivates

Definition 9.2.2. A function h that is bounded and uniformly continuous on $\mathbb{R}^d \times [0, \infty)$ is called a viscosity solution of (9.2.5) if $h(x, 0) = h_0(x)$ for all $x \in \mathbb{R}^d$ (where h_0 is also assumed to be bounded and uniformly continuous) if whenever for a smooth function η, $h - \eta$ has a local maximum (minimum) at $(x_0, t_0) \in \mathbb{R}^d \times (0, \infty)$, then

$$\frac{\partial \eta(x_0, t_0)}{\partial t} + J\left(\frac{\partial \eta(x_0, t_0)}{\partial x}, \quad x\right) \leq (\geq) 0. \tag{9.2.12}$$

In particular, the solution concept of this definition does not require any differentiability of h. For the test functions η, we can actually require that $h - \eta \leq (\geq) 0$ and $h(x_0, t_0) - \eta(x_0, t_0) = 0$. Derivatives then are only evaluated for test functions that touch h at the point in question, but not for h itself.

First of all, this solution concept is consistent in the sense that when a viscosity h is smooth, it is a classical solution of (9.2.5). This is trivial; as in that case, we may use the test function $\eta = h$ so that $h - \eta \equiv 0$ has both a local maximum and minimum at any point, and the two inequalities in (9.2.12) then yield (9.2.5). With a little more work, one shows that when a viscosity h is only known to be of class C^1 then it already is a classical solution of (9.2.5). Also, when J satisfies a Lipschitz condition, then viscosity solutions are unique. We refer to [7] for details.

The existence question is more subtle. One can use the general theory of parabolic equations to obtain the existence of a solution of (9.2.6) for any $\epsilon > 0$, as well as suitable uniform estimates that are independent of ϵ and that can be used to obtain the uniform convergence of some subsequence of h^ϵ to some function h for $\epsilon \to 0$. As we have explained, the viscosity inequalities pertain to the limit, and h therefore is a viscosity solution in the sense of the definition.

The fundamental points in the scheme are that the higher order term $\epsilon \Delta$ has a regularizing effect and that the qualitative control of the solutions gained from the maximum principle is independent of $\epsilon > 0$ and can therefore passed on to the limit for $\epsilon \to 0$.