1 ODEs and Connections to Evolution Equations

The main purpose of these notes is the study of two classes of nonlinear partial differential equations (PDEs) arising from physics: *wave equations*, and *transport equations* arising from kinetic theory (e.g., the *Vlasov-Poisson* and *Vlasov-Maxwell* equations). Both of these are subclasses of *evolution equations*, that is, PDEs that model a system evolving with respect to a "time" parameter.

When solving such evolution equations, the appropriate formulation of the problem is usually as an *initial value*, or *Cauchy, problem.* More specifically, we are given certain *initial data*, representing the state of the system at some initial time. The goal, then, is to "predict the future", that is, to find the solution of the PDE, which represents the behaviour of the system at all times.[1] Three classical examples of (linear) evolution equations are:

1. *Heat equation*: $\partial_t u - \Delta_x u = 0$, where the unknown is a function $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$. Here, the initial data is the value of u at $t = 0$, i.e., $u|_{t=0} = u_0 : \mathbb{R}^n \to \mathbb{R}$.

2. *Schrödinger equation*: $i\partial_t u + \Delta_x u = 0$, where the unknown is $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{C}$. The initial data is again the value of u at $t = 0$, i.e., $u|_{t=0} = u_0 : \mathbb{R}^n \to \mathbb{C}$.

3. *Wave equation*: $-\partial_t^2 u + \Delta_x u = 0$, where the unknown is $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$. Here, we require initial values for both u and $\partial_t u$, i.e., $u|_{t=0} = u_0$ and $\partial_t u|_{t=0} = u_1$.

In all three examples, there is one "time" dimension, denoted by $t \in \mathbb{R}$, and n "space" dimensions, denoted by $x \in \mathbb{R}^n$. Moreover, Δ_x denotes the Laplacian in the spatial variables,

$$\Delta_x := \sum_{k=1}^{n} \partial_{x^k}^2 .$$

For various technical reasons, the study of evolution equations can become quite complicated. Thus, it is beneficial to first look at some "model problems", which are technically simpler than our actual equations of interest, but still demonstrate many of the same fundamental features. A particularly useful model setting to consider is the theory of (first-order) ordinary differential equations (ODEs). The advantage of this is twofold: not only can many phenomena in evolutionary equations be demonstrated in the ODE setting, but also most readers will already have had some familiarity with ODEs.

Thus, in this section, we discuss various key aspects in the study of ODEs, and we highlight how these aspects are connected to the study of evolutionary PDE.[2]

1.1 Existence of Solutions

Throughout most of the upcoming discussion, we will consider the initial value problem for the following system of ODEs:

$$x' = y(t, x), \qquad x(t_0) = x_0. \tag{1.1}$$

Here, t is the independent variable, x is an $\mathbb{R}^n$-valued function to be solved, and the given function $y : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ defines the differential equation.[3] Recall the following:

[1] When solving on a finite domain, one requires in addition appropriate boundary conditions.
[2] A large portion of this chapter was inspired by the first chapter of T. Tao's monograph, [Tao2006].
[3] One can also restrict the domain of y to an open subset of $\mathbb{R} \times \mathbb{R}^n$, but we avoid this for simplicity.

Definition 1.1. *A differentiable function $x : I \to \mathbb{R}^n$, where I is a subinterval of $\mathbb{R}$ containing t_0, is a* solution *of* (1.1) *iff $x(t_0) = x_0$, and $x'(t) = y(t, x(t))$ for all $t \in I$.*

Such a solution x is called global *iff $I = \mathbb{R}$, and* local *otherwise.*

Abstractly, we can think of this as an evolution problem, with t in (1.1) functioning as the "time" parameter. A solution x of (1.1) can then be seen as a curve in the finite-dimensional space $\mathbb{R}^n$, parametrised by this time.

This perspective is also pertinent to evolution equations. For the sake of discussion, consider the ($(n+1)$-dimensional free) *Schrödinger equation*

$$i\partial_t u + \Delta_x u = 0 \tag{1.2}$$

where $u = u(t, x)$ is a complex-valued function of both time $t \in \mathbb{R}$ and space $x \in \mathbb{R}^n$. While the most apparent definition of a solution of (1.2) is as a sufficiently differentiable map $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{C}$, one could alternatively think of u as mapping each time t to a function $u(t)$ of n space variables. In other words, analogous to the ODE situation, one can think of a solution $t \mapsto u(t)$ as a curve in some *infinite-dimensional* space H of functions $\mathbb{R}^n \to \mathbb{C}$.

Remark 1.2. In contrast to the finite-dimensional ODE setting, where $\mathbb{R}^n$ is essentially the only appropriate space to consider, there are different possibilities one can potentially take for the infinite-dimensional space H. The choice of an appropriate H in which solutions live is one of the many challenges in solving and understanding solutions of PDEs.[4]

As we shall see below, this viewpoint of evolutionary PDEs as ODEs in an infinite-dimensional space will prove to be immediately useful. For instance, we can solve many such nonlinear PDEs using essentially the same techniques (Picard iteration, contraction mapping theorem) as for ODEs; we briefly review this existence theory in this subsection. In the remaining subsections, we discuss several other important concepts in ODEs that have direct analogues in evolutionary PDEs; examples include unconditional uniqueness arguments, Duhamel's principle, and "bootstrap" arguments for treating nonlinear terms.

The first crucial ingredient in the existence theory for ODEs is expressing (1.1) as an equivalent *integral equation.* Formally, by integrating (1.1) in t, we obtain the relation

$$x(t) = x(t_0) + \int_{t_0}^{t} y(s, x(s))ds. \tag{1.3}$$

Thus, in order to solve (1.1), it suffices to solve (1.3) instead.

Remark 1.3. One technical point to note is that one requires x to be differentiable to make sense of (1.1), while no such requirement is required for (1.3). However, for any x that satisfying the integral equation, the right-hand side of (1.3) automatically implies that x is differentiable. Thus, (1.1) and (1.3) are equivalent conditions.

On the other hand, for evolutionary PDEs, the analogous differential and integral equations will no longer be equivalent; in particular, the latter equation is often a strictly weaker requirement than the former. As a result, one must distinguish between solutions to the differential and integral equations, i.e., *classical* and *strong* solutions, respectively.

Next, we define the map Φ as follows: for an $\mathbb{R}^n$-valued curve x, we let $\Phi(x)$ be the $\mathbb{R}^n$-valued curve defined by the right-hand side of (1.3). From this viewpoint, solving (1.3)

[4] Common examples of H include $L^p(\mathbb{R}^n)$, as well as various Sobolev and Hölder spaces.

is equivalent to finding a *fixed point* of Φ, i.e., x such that

$$\Phi(x) = x. \tag{1.4}$$

To find such a fixed point, we resort to the following abstract theorem:

Theorem 1.4 (Contraction mapping theorem). *Let (X, d) be a nonempty complete metric space, and let $\Phi : X \to X$ be a* contraction, *i.e., there is some $c \in (0, 1)$ such that*

$$d(\Phi(x), \Phi(y)) \leq c \cdot d(x, y), \qquad x, y \in X. \tag{1.5}$$

Then, Φ has a unique fixed point in X.

Sketch of proof. Let x_0 be any element of X, and define the sequence (x_n) inductively by $x_{n+1} := \Phi(x_n)$. The contraction property (1.5) implies that (x_n) is a Cauchy sequence and hence has a limit x_∞. Since (1.5) also implies Φ is continuous, then

$$x_\infty = \lim_n x_{n+1} = \lim_n \Phi(x_n) = \Phi(x_\infty),$$

i.e., x_∞ is a fixed point of Φ.

For uniqueness, suppose $x, y \in X$ are fixed points of Φ. Then, (1.5) implies

$$d(x, y) = d(\Phi(x), \Phi(y)) \leq c \cdot d(x, y).$$

Since $c < 1$, it follows that $d(x, y) = 0$. □

The strategy to solving (1.4) is to show that this Φ is indeed a contraction on the appropriate space. Then, Theorem 1.4 yields a fixed point of Φ, which is the solution of (1.1). The precise result is stated in the subsequent theorem:

Theorem 1.5 (Existence of solutions). *Consider the initial value problem* (1.1)*, and let $\Omega_{\mathcal{T},\mathcal{R}}$, where $\mathcal{T}, \mathcal{R} > 0$, be the following closed neighbourhood:*

$$\Omega_{\mathcal{T},\mathcal{R}} := \{(t, x) \mid |t - t_0| \leq \mathcal{T},\ |x| \leq 2\mathcal{R}\}.$$

Suppose also that the function y in (1.1) *satisfies the following:*

- *y is uniformly bounded on $\Omega_{\mathcal{T},\mathcal{R}}$—there exists $M > 0$ such that*

$$|y(t, x)| \leq M, \qquad (t, x) \in \Omega_{\mathcal{T},\mathcal{R}}. \tag{1.6}$$

- *y satisfies the following Lipschitz property on $\Omega_{\mathcal{T},\mathcal{R}}$—there exists $L > 0$ such that*

$$|y(t, x_1) - y(t, x_2)| \leq L|x_1 - x_2|, \qquad (t, x_1), (t, x_2) \in \Omega_{\mathcal{T},\mathcal{R}}. \tag{1.7}$$

Then, given any $x_0 \in \mathbb{R}^n$ with $|x_0| \leq \mathcal{R}$, the initial value problem (1.1) *has a solution $x : [t_0 - T, t_0 + T] \to \mathbb{R}^n$, for some $T \in (0, \mathcal{T})$ depending on $\mathcal{R}$, M, and L.*

Remark 1.6. Note the time T of existence of the solution in Theorem 1.5 depends on the size $|x_0|$ of the initial data. This is manifested in the dependence of T on $\mathcal{R}$.

Proof. Let $T > 0$, whose value will be determined later, and set $I = [t_0 - T, t_0 + T]$. Consider first the space $C(I; \mathbb{R}^n)$ of all continuous functions $x : I \to \mathbb{R}^n$, along with the sup-norm

$$\|x\| := \sup_{t \in I} |x(t)|.$$

In particular, the above forms a Banach space. Consider next the closed ball

$$X := \{x \in C(I;\mathbb{R}^n) \mid \|x\| \leq 2\mathcal{R}\}.$$

Then, X, being closed, forms a complete metric space along with the induced metric

$$d(x_1, x_2) = \|x_2 - x_1\|.$$

Now, we define Φ to be the map arising from the integral equation (1.3):

$$\Phi : C(I;\mathbb{R}^n) \to C(I;\mathbb{R}^n), \qquad [\Phi(x)](t) := x_0 + \int_{t_0}^{t} y(s, x(s)) ds.$$

Then, in order to generate a fixed point of Φ, and hence a solution of (1.1), we must show:

1. Φ maps X into X.
2. $\Phi : X \to X$ is a contraction.

For the first point, given $x \in X$, we see from the definition of Φ that

$$\|\Phi(x)\| \leq |x_0| + \int_I |y(s, x(s))| ds \leq \mathcal{R} + 2MT.$$

Then, for small enough T (depending on M), we have that $\|\Phi(x)\| \leq 2\mathcal{R}$, i.e., $\Phi(x) \in X$. This proves that Φ indeed maps X into X. For the second point, given $x_1, x_2 \in X$,

$$\begin{aligned} \|\Phi(x_2) - \Phi(x_1)\| &\leq \int_I |y(s, x_2(s)) - y(s, x_1(s))| ds \\ &\leq L \int_I |x_2(s) - x_1(s)| ds \\ &\leq 2TL \|x_2 - x_1\|. \end{aligned}$$

Taking T small enough so that $TL < \frac{1}{4}$, then Φ is a contraction, completing the proof. □

Remark 1.7. Note that the proof the contraction mapping theorem is both relatively simple and more intuitive than the theorem statement itself. Thus, a more direct approach to proving Theorem 1.5 can be achieved by running the proof of the contraction mapping theorem directly for the Φ in (1.4), rather than applying the abstract Theorem 1.4. This process, which amounts to constructing an infinite sequence of closer and closer approximations to the desired solution of (1.1), is known as *Picard iteration*.

In the PDE setting, for instance for a nonlinear Schrödinger or wave equation such as

$$i\partial_t u + \Delta_x u = \pm |u|^2 u, \qquad -\partial_t^2 u + \Delta_x u = \pm |u|^2 u,$$

the main ideas for proving existence of solutions are largely the same as in Theorem 1.5. Indeed, the first step is to write the above differential equations as integral equations, which can again be cast as a fixed point problem. The goal then is once again to show that the map Φ obtained from the integral equation is a contraction.

In this setting, Φ now acts on a space of curves mapping into an infinite-dimensional space H of functions on $\mathbb{R}^n$. One of the main challenges, then, is to choose an appropriate space H with an appropriate norm so that Φ can be shown to be contraction. What H can be will of course depend very crucially on the PDE under consideration.

1.2 Uniqueness of Solutions

If an ODE models some system in the real world, then solving an initial value problem roughly amounts to "predicting the future", given the initial state of the system. However, this is not entirely accurate—although Theorem 1.5 states that solutions *exist* (for nice enough ODE), we have not discussed whether such solutions are *unique*. If multiple solutions were to exist, then the problem—at least in the way that we have stated it—is not deterministic, and we cannot "predict the future", so to speak.

The attentive reader may have noticed that the contraction mapping theorem, applied in the proof Theorem 1.5, guaranteed the existence of a unique fixed point. While this seems at first glance to take care of the uniqueness problem, there is, unfortunately, a hole in this chain of reasoning. Indeed, *the contraction mapping theorem only guarantees that the fixed point, or solution, is unique on the space* X, which is a closed ball of the larger Banach space $C([t_0 - T, t_0 + T]; \mathbb{R}^n)$. In particular, this does not rule out a different solution $z \in C([t_0 - T, t_0 + T]; \mathbb{R}^n)$ that grows large, i.e., that does not lie within the ball X.

The uniqueness obtained from the contraction mapping theorem is commonly referred to as *conditional uniqueness*—in the context of Theorem 1.5, we know that a solution x of (1.1) is unique as long as $\|x\| \leq 2\mathcal{R}$, that is, as x does not grow too large. The statement that we want, however, is *unconditional uniqueness*—that the solution x of (1.1) is in fact the only solution of all sizes, i.e., it is unique in all of $C([t_0 - T, t_0 + T]; \mathbb{R}^n)$.

Theorem 1.8 (Uniqueness of solutions). *Consider the initial value problem* (1.1)*. Suppose also that* y *in* (1.1) *satisfies the following locally Lipschitz property: for each*

$$\Omega_{\mathcal{T},\mathcal{R}} := \{(t, x) \in \mathbb{R} \times \mathbb{R}^n \mid |t - t_0| \leq \mathcal{T},\ |x| \leq 2\mathcal{R}\}, \qquad \mathcal{T}, \mathcal{R} > 0,$$

there exists some $L_{\mathcal{T},\mathcal{R}} > 0$ *such that*

$$|y(t, x_1) - y(t, x_2)| \leq L_{\mathcal{T},\mathcal{R}} |x_1 - x_2|, \qquad (t, x_1), (t, x_2) \in \Omega_{\mathcal{T},\mathcal{R}}. \tag{1.8}$$

Let $T > 0$*, and assume* $x_1, x_2 \in C([t_0 - T, t_0 + T]; \mathbb{R}^n)$ *are two solutions to* (1.1)*. Then,* $x_1(t) = x_2(t)$ *for all* $t \in [t_0 - T, t_0 + T]$.

Remark 1.9. Notice the Lipschitz condition in Theorem 1.8 is analogous to that in Theorem 1.5. The slight difference arises from the fact that here, one must assume y remains "nice", in the sense of Theorem 1.5, no matter how large the solutions x_i may get.

To obtain such an unconditional uniqueness statement, one generally requires another argument in addition to the proof of existence. For the most part, such uniqueness arguments are relatively simple, in that they use similar tools as the existence arguments.[5]

Theorem 1.8 can be proved in multiple ways. One of the most straightforward is via a linear estimate known as *Gronwall's inequality*. In fact, Gronwall's inequality is also an incredibly useful tool in the study of PDEs, for similar unconditional uniqueness arguments as well as other applications. The main idea derives from the method of integrating factors used in basic ODE theory, along with the observation that they are applicable to *inequalities* as well as to equations. We present some special cases here:

Theorem 1.10 (Gronwall inequality). *Let* $x, C : [0, T] \to [0, \infty)$.

[5] In some PDEs settings, unconditional uniqueness arguments can sometimes be much more nontrivial. Indeed, many such statements have only recently been proved, or even remain open.

1. Differential version: *Assume x is differentiable, and x satisfies*

$$x'(t) \leq C(t) \cdot x(t), \qquad t \in [0, T]. \tag{1.9}$$

Then, x also satisfies

$$x(t) \leq x(0) \cdot \exp\left[\int_0^t C(s)ds\right]. \tag{1.10}$$

2. Integral version: *Assume x is continuous, and x satisfies*

$$x(t) \leq A(t) + \int_0^t C(s)x(s)ds, \qquad t \in [0, T]. \tag{1.11}$$

for some nondecreasing $A : [0, T] \to [0, \infty)$. Then, x also satisfies

$$x(t) \leq A(t) \cdot \exp\left[\int_0^t C(s)ds\right]. \tag{1.12}$$

Proof. For the differential version, we multiply (1.9) by $\exp[-\int_0^t C(s)ds]$, which yields

$$\frac{d}{dt}[e^{-\int_0^t C(s)ds} x(t)] \leq 0, \qquad t \in [0, T].$$

Integrating the above from 0 to t results in (1.10).

For the integral version, we define

$$z(t) := \exp\left[-\int_0^t C(s)ds\right] \int_0^t C(s)x(s)ds.$$

Differentiating z, we see that

$$\begin{aligned} z'(t) &= C(t) \exp\left[-\int_0^t C(s)ds\right] \left[x(t) - \int_0^t C(s)x(s)ds\right] \\ &\leq A(t)C(t) \exp\left[-\int_0^t C(s)ds\right]. \end{aligned}$$

Since $z(0) = 0$ and A is nondecreasing, we have

$$\begin{aligned} z(t) &\leq \int_0^t A(s)C(s) \exp\left[-\int_0^s C(r)dr\right] ds \\ &\leq A(t) \int_0^t C(s) \exp\left[-\int_0^s C(r)dr\right] ds \\ &= A(t) - A(t) \exp\left[-\int_0^t C(s)\right]. \end{aligned}$$

Finally, by (1.11), we obtain, as desired,

$$x(t) \leq A(t) + \exp\left[\int_0^t C(s)ds\right] \cdot z(t) \leq A(t) \cdot \exp\left[\int_0^t C(s)ds\right]. \qquad \square$$

We now apply Gronwall's inequality to prove Theorem 1.8.

Proof of Theorem 1.8. Let us assume for convenience that $t_0 = 0$. Since both x_1 and x_2 are bounded on $[0, 0 + T]$, then (1.8) implies for any $t \in [0, T]$ that

$$|x_1(t) - x_2(t)| \leq \int_0^t |y(s, x_1(s)) - y(s, x_2(s))|ds \leq L \int_0^t |x_1(s) - x_2(s)|ds.$$

Applying the integral Gronwall's inequality, (1.12), with $x := |x_1 - x_2|$ and $A :\equiv 0$ yields

$$|x_1(t) - x_2(t)| \leq 0 \cdot \exp t = 0, \qquad 0 \leq t \leq T.$$

An analogous argument also shows that $x_1 = x_2$ on $[-T, 0]$. $\square$

Now that both existence and uniqueness have been established, one can discuss what is the largest time interval that a solution exists. The basic argument is as follows. First, Theorem 1.5 furnishes a (unique) solution x on $[t_0 - T, t_0 + T]$. Next, one can solve the same ODE, but with initial data given by x at times $t_0 \pm T$. Theorem 1.8 guarantees that these new solutions coincide with x wherever both are defined.

Thus, "patching" together these solutions yields a new solution x of (1.1) on a larger interval $[t_0 - T_1, t_0 + T_2]$. By iterating this process indefinitely, we obtain:

Corollary 1.11 (Maximal solutions). *Consider the initial value problem* (1.1)*, and suppose y satisfies the same hypotheses as in Theorem 1.8. Then, there exists a "maximal" interval (T_-, T_+) containing t_0, where $-\infty \leq T_- < T_+ \leq \infty$, such that:*

- *There exists a solution $x : (T_-, T_+) \to \mathbb{R}^n$ to* (1.1)*.*
- *x is the only solution to* (1.1) *on the interval (T_-, T_+).*
- *If $\tilde{x} : I \to \mathbb{R}^n$ is another solution of* (1.1)*, then $I \subseteq (T_-, T_+)$.*

We refer to x in Corollary 1.11 as the *maximal solution* of (1.1). In fact, one can say a bit more about the behaviour of the maximal solutions at the boundaries $T_\pm$.

Corollary 1.12 (Breakdown criterion). *Consider the initial value problem* (1.1)*, and suppose y satisfies the same hypotheses as in Theorem 1.8. Let $x : (T_-, T_+) \to \mathbb{R}^n$ be the maximal solution of* (1.1)*. Then, if $T_+ < \infty$, then*

$$\limsup_{t \nearrow T_+} |x(t)| = \infty. \tag{1.13}$$

An analogous result holds at T_-.

Proof. If (1.13) fails to hold, then $|x|$ is uniformly bounded near T_+. Since the time of existence in Theorem 1.5 depends only on the size of the initial data, then one can solve the ODE with initial data at a time $T_+ - \varepsilon$ for arbitrarily small $\varepsilon > 0$, but always for a fixed amount of time. By uniqueness, this allows us to push the solution past time T_+, contradicting that x is the maximal solution. ☐

Finally, we remark that in the setting of nonlinear evolution equations, one can often use Gronwall's inequality in a similar fashion as in Theorem 1.8 in order to show unconditional uniqueness. Then, the same argument behind Corollary 1.11 yields an analogous notion of maximal solutions for PDEs. Furthermore, for a large subclass of such PDEs—known as "subcritical"—one can establish that the time of existence of solutions depend only on the size of the initial data.[6] As a result, an analogue of Corollary 1.12 holds for these PDEs.

Later, we will formally demonstrate all these points for nonlinear wave equations.

1.3 Duhamel's Principle

We turn our attention to linear systems of ODE. Consider first the homogeneous case,

$$x' = y(t, x) = Ax, \qquad x(t_0) = x_0 \in \mathbb{R}^n, \tag{1.14}$$

where A is a constant $n \times n$ matrix.

[6] Here, "size" is measured by an appropriate norm on the infinite-dimensional space H of functions.

When $n = 1$, then A can be expressed as a constant $\lambda \in \mathbb{R}$. The resulting system $x' = \lambda x$ then has an explicit solution, $x(t) = e^{(t-t_0)\lambda} x_0$, which, depending on the sign of λ, either grows exponentially, decays exponentially, or stays constant.

In higher dimensions, one can still write the solution in the same manner

$$x(t) = e^{tA} x_0, \tag{1.15}$$

where e^{tA} is the matrix exponential, which, for instance, can be defined via a Taylor series:

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}.$$

In (1.15), the matrix e^{tA} is multiplied to x_0, represented as a column vector. The operator $x_0 \mapsto e^{tA} x_0$ is often called the *linear propagator* of the equation $x' = Ax$.

To better understand the solution $e^{tA} x_0$, one generally works with a basis of $\mathbb{R}^n$ which diagonalises A (or at least achieves Jordan normal form). In particular, by considering the eigenvalues of A, one can separate the solution curve x into individual directions which grow exponentially, decay exponentially, or oscillate.

Remark 1.13. In the more general case in which A also depends on t, one can still define linear propagators, though they may not be representable as matrix exponentials.

This reasoning extends almost directly to the PDE setting as well. Consider for instance the initial value problem for the free linear Schrdinger equation, which can be written as

$$\partial_t u = i \Delta_x u, \qquad u|_{t=0} = u_0 : \mathbb{R}^n \to \mathbb{C}.$$

Formally at least, thinking of $i\Delta_x$ as the (constant in time) linear operator on our infinite-dimensional space of functions, we can write the solution of the initial value problem in terms of a linear propagator,[7]

$$u(t, x) = (e^{it\Delta_x} u_0)(x).$$

Similarly, for a free linear transport equation,

$$\partial_t u + v \cdot \nabla_x u = 0, \qquad u|_{t=0} = u_0 : \mathbb{R}^n \to \mathbb{R},$$

where $v \in \mathbb{R}^n$, one can write a similar linear propagator,

$$u = e^{-t(v \cdot \nabla_x)} u_0,$$

although in this case the solution has a simpler explicit formula,

$$u(t, x) = u_0(x - tv).$$

Later, we will study the propagator for the wave equation in much greater detail.

Next, we consider inhomogeneous linear systems containing a forcing term,

$$x' = Ax + F, \qquad x(t_0) = x_0, \tag{1.16}$$

[7] There are multiple ways to make precise sense of the operators $e^{it\Delta_x}$. For example, this can be done using Fourier transforms, or through techniques from spectral theory.

where A is as before, and $F : \mathbb{R} \to \mathbb{R}^n$. To solve this system, one can apply the matrix analogue of the method of integrating factors from ODE theory. In particular, multiplying (1.16) by e^{-tA}, we can rewrite it as

$$(e^{-tA}x)' = e^{-tA}F.$$

Integrating the above with respect to t yields:

Proposition 1.14. *The solution to* (1.16) *is given by*

$$x(t) = e^{tA}x_0 + \int_{t_0}^{t} e^{(t-s)A}F(s)ds. \tag{1.17}$$

The first term $e^{tA}x_0$ in (1.17) is the solution to the homogeneous problem (1.14), while the other term represents the solution to the inhomogeneous equation with zero initial data. In the case that F is "small", one can think of (1.17) as the solution $e^{tA}x_0$ of the homogeneous problem plus a perturbative term. This should be contrasted with (1.3), which expresses the solution as a perturbation of the constant curve.

This viewpoint is especially pertinent for nonlinear equations. Consider the system

$$x' = -x + |x|x,$$

for (1.17) yields

$$x(t) = e^{-t}x_0 + \int_{t_0}^{t} e^{-(t-s)}[|x(s)|x(s)]ds.$$

Then, for small x or for small times $t - t_0$, the above indicates that x should behave like the linear equation, and that the nonlinear effects are perturbative.

Again, these ideas extend to PDE settings; the direct analogue of (1.17) in the PDE setting is often called *Duhamel's principle*. For example, a commonly studed family of nonlinear dispersive equations is the *nonlinear Schrödinger equation* (*NLS*), given by

$$i\partial_t u + \Delta_x u = \pm |u|^{p-1}u, \qquad p > 1.$$

Then, Duhamel's principle implies that

$$u(t) = e^{i(t-t_0)\Delta_x}u_0 \mp i \int_{t_0}^{t} e^{i(t-s)\Delta_x}[|u(s)|^{p-1}u(s)]ds.$$

In fact, in this case, the Picard iteration process is applied directly on the above formula, as it captures more effectively the qualitative properties of the solution.

As we will discuss in much more detail later, a similar Duhamel's formula exists for wave equations, and it has similar uses as for the NLS.

1.4 Continuity Arguments

The main part of these notes will deal with nonlinear PDE, for which solutions usually cannot be described by explicit equations. Thus, one must resort to other tools to capture various qualitative and quantative aspects of solutions.

One especially effective method is called the *continuity argument*, sometimes nicknamed "*bootstrapping*". The main step in this argument is to *assume what you want to prove* and then to *prove a strictly better version of what you assumed*. Until the process is properly explained, it seems suspiciously like circular reasoning. Moreover, because it is used so often in studying nonlinear evolution equations, continuity arguments often appear in research papers without comment or explanation, which can be confusing to many new readers.

Rather than discussing the most general possible result, let us consider as a somewhat general example the following "trivial" proposition:

Proposition 1.15. *Let $f : [0, T) \to [0, \infty)$ be continuous, where $0 < T \leq \infty$, and fix a constant $C > 0$. Suppose the following conditions hold:*

1. *$f(0) \leq C$.*
2. *If $f(t) \leq 4C$ for some $t > 0$, then in fact $f(t) \leq 2C$.*

Then, $f(t) \leq 4C$ (and hence $f(t) \leq 2C$) for all $t \in [0, T)$.

The intuition behind Proposition 1.15 is simple. Assumption (1) implies that f starts below $4C$. If f were to grow larger than $4C$, then it must cross the threshold $4C$ at some initial point t_0. But then, assumption (2) implies that f actually lies below $2C$, so that f could not have reached $4C$ at time t_0.

In applications of Proposition 1.15 (or some variant), the main problem is to show that assumption (2) holds. In other words, we assume what we want to prove ($f(t) \leq 4C$, called the *bootstrap assumption*), and we prove something strictly better ($f(t) \leq 2C$).

For completeness, let us give a more robust topological proof of Proposition 1.15:

Proof. Let $A := \{t \in [0, T) \mid f(s) \leq 4C \text{ for all } 0 \leq s \leq t\}$. Note that A is nonempty, since $0 \in A$, and that A is closed, since f is continuous. Now, if $t \in A$, then the second assumption implies $f(t) \leq 2C$, so that $t + \delta \in A$ for small enough $\delta > 0$. Thus, A is a nonempty, closed, and open subset of the connected set $[0, T)$, and hence $A = [0, T)$. □

In either the ODE or the PDE setting, one can think of $f(t)$ as representing the some notion of "size" of the solution up to time t. Of course, in general, one cannot compute explicitly how large $f(t)$ is. However, in order to better understand the behaviour of, or to further extend (say, using Corollary 1.12) the lifespan of, solutions, one often wishes to prove bounds on $f(t)$.[8] Continuity arguments, for instance via Proposition 1.15, provide a method for achieving precisely this goal, without requiring explicit formulas for the solution.

To see bootstrapping in action, let us consider two (ODE) examples:

Example 1.16. Let $n = 1$, and consider the nonlinear system

$$x'(t) = \frac{|x(t)|^2}{1 + t^2}, \qquad x(0) = x_0 \in \mathbb{R}. \tag{1.18}$$

We wish to show the following: *If $|x_0|$ is sufficiently small, then the solution to* (1.18) *is global, i.e., $x(t)$ is defined for all $t \in \mathbb{R}$. Furthermore, x is everywhere uniformly small.*

By the breakdown criterion, Corollary 1.12, it suffices to prove the appropriate uniform bound for x, since this implies x can be further extended.[9] For this, let

$$f(t) = \sup_{0 \leq s \leq t} |x(s)|.$$

[8]In fact, in many PDE settings, estimates for this $f(t)$ are often essential to *solving* the equation itself.

[9]More precisely, one lets (T_-, T_+) be the maximal interval of definition for x, and one applies the continuity argument to show that x is small on this domain. Corollary 1.12 then implies that $T_\pm = \pm\infty$.

Moreover, let $\varepsilon > 0$ be a small constant, to be determined later, and suppose $|x_0| \leq \varepsilon$.

For the continuity argument, let us impose the bootstrap assumption

$$f(T) = \sup_{0 \leq t \leq T} |x(t)| \leq 4\varepsilon, \qquad T > 0.$$

Then, for any $0 \leq t \leq T$, we have, from (1.3) and the bootstrap assumption,

$$|x(t)| \leq |x_0| + \int_0^t \frac{|x(s)|^2}{1+s^2} ds \leq \varepsilon + 16\varepsilon^2 \int_0^t \frac{1}{1+s^2} ds.$$

Since $t \mapsto (1+t^2)^{-1}$ is integrable on $[0, \infty)$, then

$$|x(t)| \leq \varepsilon + C\varepsilon^2,$$

and as long as ε is sufficiently small (with respect to C), we have $|x(t)| \leq 2\varepsilon$, and hence $f(T) \leq 2\varepsilon$ (a strictly better result). Proposition 1.15 now implies $|x(t)| \leq 2\varepsilon$ for all $t \geq 0$.

An analogous argument proves the same bound for $t \leq 0$, hence x is small for all times.

Example 1.17. Let x be a solution of (1.1), and suppose x satisfies

$$|x| \leq A + B|x|^p, \qquad A, B > 0, \quad 0 < p < 1.$$

We wish to show that x *is uniformly bounded* and that x *is a global solution.*

Again, by Corollary 1.12, we need only show a uniform bound for x. Let

$$f(t) = \sup_{0 \leq s \leq t} |x(s)|,$$

and assume $f(T) \leq 4C$ for some sufficiently large C. Then, by our bootstrap assumption,

$$|x(t)| \leq A + B(4C)^p \leq A + 4^p BC^p, \qquad 0 \leq t \leq T$$

which for C large enough implies $f(T) \leq 2C$.

By Proposition 1.15, the above shows that x is uniformly bounded for all positive times. An analogous argument controls negative times as well.

The above examples give simple analogues for how continuity arguments can be used to study nonlinear PDE. Such arguments will be essential later once we consider global existence questions for nonlinear wave equations.