

# Math 201 Handout 1

Rich Schwartz

September 6, 2006

The purpose of this handout is to give a proof of the basic existence and uniqueness result for ordinary differential equations. This result includes the statement that the solutions depend smoothly on the initial conditions if the data for the differential equation is smooth. (See below for a precise formulation.) There are plenty of accounts of this theorem in the literature. However, I found the various written accounts somehow complicated and opaque. I hope to give a streamlined and clear proof here. For ease of exposition, I will give the proof in  $\mathbf{R}^2$ . This case shares all the features of the general case of  $\mathbf{R}^n$ . Mainly, I'm doing  $\mathbf{R}^2$  instead of  $\mathbf{R}^n$  so that I can write  $(x, y)$  in place of  $(x_1, \dots, x_n)$ .

## 1 The Result

We will always coordinatize  $\mathbf{R}^3$  as  $(x, y, t)$ . We think of  $t$  as “time”. A map  $\phi : \mathbf{R}^3 \rightarrow \mathbf{R}$  is *smooth* if all orders of partial derivatives of  $\phi$  exist and are continuous. (Just the existence of all orders of derivatives forces their continuity.) I'll denote partial derivatives by  $d_t, d_x, d_y$ . For instance  $d_t\phi = \partial\phi/\partial t$ .

Suppose that  $V : U \rightarrow \mathbf{R}^2$  is a smooth function defined in a neighborhood  $U$  of  $(0, 0)$ . We can write  $V = (V_1, V_2)$ . Given a map  $F : U \rightarrow \mathbf{R}^2$  we will write  $F = (F_1, F_2)$ . Here  $F_1$  and  $F_2$  are both maps from  $\mathbf{R}^2 \rightarrow \mathbf{R}$ . The basic ordinary differential equation is given by

$$d_t F = V \circ F.$$

Geometrically,  $V$  defines a vector field in a neighborhood of the origin and the curve  $\gamma(t) = F(x, y, t)$  (for  $x$  and  $y$  held fixed) is always tangent to  $V$ .

To write out the above equation in detail, we have

- $d_t F_1(x, y) = V_1(F_1(x, y), F_1(x, y))$ .
- $d_t F_2(x, y) = V_2(F_2(x, y), F_2(x, y))$ .

Assuming that  $F$  exists, we define

$$f(x, y) = F(x, y, 0).$$

We say that the function  $F$  is a  $V$ -extension of  $f$  in this case. In particular,  $F$  should be defined both for small positive values of  $t$  and small negative values.

Here is the basic result:

**Theorem 1.1** *Suppose that  $U$  is a small neighborhood of the origin in  $\mathbf{R}^2$  on which  $V$  is defined and smooth. Suppose that  $f : U \rightarrow \mathbf{R}^2$  is a smooth function. Then  $f$  has a smooth  $V$ -extension  $F$ , defined in perhaps a smaller neighborhood about the origin.*

## 2 Existence and Uniqueness

Here we will hold the point  $(x, y)$  fixed. We show the existence and uniqueness of a smooth map  $\phi(t) = F(x, y, t)$  which satisfies the basic equation, at least in some interval  $(-\epsilon, \epsilon)$ . The interval  $(-\epsilon, \epsilon)$  only depends on some properties of  $V$  which are the same for all choices  $(x, y)$ . Thus, when we run our argument at each point  $(x, y)$  in a neighborhood of  $(0, 0)$ , we can use the same value of  $\epsilon$  for all points. This gives us our  $V$ -extension  $F$  of  $f$ . All we know about  $F$  from this construction is that the partial derivatives  $d_t^{(n)} F$  exist for all  $n$ . We don't know about the other partial derivatives. Dealing with these other partial derivatives, which we do in later sections, is the most painful part of the argument.

So, suppose now that the closed interval  $I = [-\epsilon, \epsilon]$  is fixed. We will specify the choice of  $\epsilon$  later on. We let  $X$  denote the set of all continuous functions  $\psi$  from  $I$  to  $\mathbf{R}^2$  such that  $\psi(0) = F(0, 0, 0)$ . Our desired solution  $\phi$  is going to be an element of  $X$ . There is a natural metric on  $X$ , namely

$$d(g_1, g_2) = \sup_{t \in I} \|g_1(t) - g_2(t)\|.$$

Here  $\|\cdot\|$  is just the Euclidean norm on  $\mathbf{R}^2$ . It is not hard to show that  $X$  is a complete metric space. That is, any Cauchy sequence in  $X$  converges to a

unique point in  $X$ . This result is sometimes stated in terms of *equicontinuous families* of functions in a basic analysis course.

A map  $T : X \rightarrow X$  is called a *contraction* if there is some  $c \in (0, 1)$  such that

$$d(T(g_1), T(g_2)) \leq cd(g_1, g_2)$$

If  $T$  is a contraction, then  $T$  has a unique fixed point. The uniqueness is immediate. For existence, you just look at the sequence  $\{T^n(x)\}$ , where  $x \in X$  is any starting point, and note that this is a Cauchy sequence.

Now for our proof. Consider the following map  $T : X \rightarrow X$ . Given any function  $\psi : [-\epsilon, \epsilon] \rightarrow \mathbf{R}$  we define

$$T\psi(s) = \psi(0) + \int_0^s V(\psi(t)) dt.$$

**Lemma 2.1** *If  $\epsilon$  is small enough then  $T$  is a contraction.*

**Proof:** Since  $V$  is smooth in a neighborhood of the origin, we have a bound  $K$  on all the first order directional derivatives of  $V$  in a possibly smaller neighborhood of the origin. This gives us a bound

$$\|V(p_1) - V(p_2)\| \leq K\|p_1 - p_2\|,$$

for any two points  $p_1$  and  $p_2$  sufficiently close to the origin. (In the literature, this condition is summarized by saying the  $V$  is *Lipschitz* near the origin.)

Taking  $\epsilon = K/2$  we have

$$\begin{aligned} \|T\psi_1(s) - T\psi_2(s)\| &\leq \int_0^s K\|\psi_1(t) - \psi_2(t)\| dt \leq K|s| d(\psi_1, \psi_2) \\ &\leq K\epsilon d(\psi_1, \psi_2) \leq d(\psi_1, \psi_2)/2. \end{aligned}$$

Since this inequality holds for all  $s \in I$ , we have

$$d(T\psi_1, T\psi_2) \leq d(\psi_1, \psi_2)/2.$$

Hence  $T$  is a contraction. ♠

So, we can start with any function, say the constant function  $\psi$ . Then the limit  $\phi = \lim T^n\psi$  exists and is the unique fixed point of  $T$ . Hence

$$\phi(s) = T\phi(s) = \phi(0) + \int_0^s V(\phi(t)) dt. \tag{1}$$

Now for the fun: We know that  $\phi$  is continuous because it is a member of  $X$ . Hence  $V \circ \phi$  is continuous. Hence  $\phi$  is the integral of a continuous function. Hence  $\phi$  is differentiable. Hence  $V \circ \phi$  is differentiable. But now  $\phi$  is the integral of a differentiable function. Hence  $\phi$  is twice differentiable. And so on. This shows inductively that  $\phi$  is smooth. When we differentiate both sides of the last equation and apply the Fundamental Theorem of Calculus, we arrive at

$$d_t \phi = V \circ \phi.$$

This proves the existence and uniqueness of solutions to the basic differential equation.

## 2.1 A General Estimate

In the last section we used the Lipschitz condition on  $V$ . That is

$$\|V(p_1) - V(p_2)\| \leq K\|p_1 - p_2\|,$$

for any two points  $p_1$  and  $p_2$  sufficiently close to the origin. Here we will prove a general estimate about two different solutions to two different equations. Suppose that  $\widehat{V}$  is some other smooth and  $K$ -Lipschitz function defined in a neighborhood of the origin.

We want to consider solutions to the two equations

$$d_t F = V \circ F; \quad d_t \widehat{F} = \widehat{V} \circ \widehat{F},$$

where  $F$  is a  $V$ -extension of  $f$  and  $\widehat{F}$  is a  $\widehat{V}$ -extension of another smooth function  $\widehat{f}$ . Intuitively, we want to prove the result that the extensions  $F$  and  $\widehat{F}$  are close provided that  $f$  and  $\widehat{f}$  are close and  $V$  and  $\widehat{V}$  are close.

**Lemma 2.2** *Suppose that  $t \in [-t_0, t_0]$  where  $\exp((K + 1)t_0) = 2$ . If*

- $\|(\widehat{x}, \widehat{y}) - (x, y)\| < \epsilon,$
- $\|\widehat{f}(\widehat{x}, \widehat{y}) - f(x, y)\| < \epsilon,$
- $\|\widehat{V}(p) - V(p)\| < \epsilon$  for all points  $p$  in a neighborhood of the origin.

then  $\|\widehat{F}(\widehat{x}, \widehat{y}, s) - F(x, y, s)\| \leq 2\epsilon.$

**Proof:** Using the integral form for the basic equation (Equation 1), and the chain rule, we have

$$\begin{aligned}
& \|\widehat{F}(\widehat{x}, \widehat{y}, \widehat{s}) - F(x, y, s)\| \\
& \leq \|f(x, y) - \widehat{f}(\widehat{x}, \widehat{y})\| + \int_0^s \|V(F(x, y, t)) - \widehat{V}(\widehat{F}(\widehat{x}, \widehat{y}, t))\| dt \\
& \leq \|f(x, y) - \widehat{f}(\widehat{x}, \widehat{y})\| + \int_0^s \|V(F(x, y, t)) - \widehat{V}(F(x, y, t))\| dt \\
& \quad + \int_0^s \|\widehat{V}(F(x, y, t)) - \widehat{V}(F(\widehat{x}, \widehat{y}, t))\| dt \\
& \leq \epsilon + \epsilon s + \int_0^s K \|F(x, y, t) - \widehat{F}(x, y, t)\| dt.
\end{aligned}$$

The last line just comes from putting in all the hypotheses of the lemma.

It is convenient to write

$$A(t) = \|F(x, y, t) - \widehat{F}(\widehat{x}, \widehat{y}, t)\|.$$

The calculation above shows that

$$A(s) \leq \epsilon + \epsilon s + \int_0^s KA(t)dt.$$

Differentiating both sides and remembering the initial conditions, we see that

$$A(0) \leq \epsilon; \quad d_t A \leq \epsilon + KA.$$

Suppose, for the sake of contradiction that there is some  $t \in [-t_0, t_0]$  such that  $A(t) > 2\epsilon$ . Without loss of generality assume  $t > 0$ . There is some maximal interval  $[t_1, t_2] \in [0, t_0]$  such that  $t \in [t_1, t_2]$  and  $A \geq \epsilon$  on all of  $[t_1, t_2]$ . For this interval we have the bound

$$A(t_1) = \epsilon; \quad d_t A \leq (K + 1)A.$$

But then

$$A(t) \leq \exp((K + 1)(t - t_0)) \times \epsilon \leq 2\epsilon,$$

which is a contradiction. ♠

As an immediate application of our result we can take  $\widehat{V} = V$  and  $\widehat{F} = F$ . The above result shows that

$$\|F(x, y, t) - F(\widehat{x}, \widehat{y}, t)\| \leq 2\|f(x, y) - f(\widehat{x}, \widehat{y})\|.$$

as long as  $t$  is sufficiently small and  $(\widehat{x}, \widehat{y})$  is sufficiently close to  $(x, y)$ . This shows that the function  $F(x, y, t)$  is continuous in all variables, sufficiently close to the origin.

### 3 Smoothness

We fix some small value of  $h$  and consider the new function

$$\widehat{F}_h(x, y, t) = \frac{F(x+h, y, t) - F(x, y, t)}{h}.$$

Of course, we might need to work in a smaller neighborhood so that everything in this last equation is defined. We also define  $\widehat{f}_h(x, y) = \widehat{F}_h(x, y, t)$ .

Our goal is to show that

$$\lim_{h \rightarrow 0} \widehat{F}_h(x, y, t)$$

exists in a neighborhood of the origin in  $\mathbf{R}^3$ . Below we will prove

**Lemma 3.1** *There is a smooth function  $\widehat{V}_h(x, y)$  such that  $\widehat{F}_h$  satisfies the differential equation*

$$\partial_t \widehat{F}_h = \widehat{V}_h \circ \widehat{F}_h.$$

*Furthermore,  $\widehat{V}_h$  is  $K$ -Lipschitz in a neighborhood of the origin, where  $K$  and the neighborhood are independent of  $h$ . Finally, the limit*

$$W = \lim_{h \rightarrow 0} \widehat{V}_h(x, y, t)$$

*exists at all points in a neighborhood of the origin in  $\mathbf{R}^3$ .*

Given Lemma 3.1 we compute the proof as follows. For any  $\epsilon > 0$  of interest to us, we choose  $h$  small enough such that

- $\|\widehat{f}_h(x, y) - d_x f(x, y)\| < \epsilon$
- $\|\widehat{V}_h(p) - W(p)\| < \epsilon$  for all  $p$  in a neighborhood  $U$  of the origin.

From the basic existence theorem we know that  $d_x f$  has a continuous  $W$ -extension  $G$ . From the estimate in the previous section we know that

$$\|\widehat{F}_h(x, y, t) - G(x, y, t)\| \leq 2\epsilon$$

throughout our neighborhood. We have  $\epsilon \rightarrow 0$  as  $h \rightarrow 0$ . Hence  $\widehat{F}_h \rightarrow G$ . Hence  $d_x F$  exists in a neighborhood of the origin. The same argument shows that  $d_y F$  exists in a neighborhood of the origin.

Before we prove Lemma 3.1 we deal with the higher derivatives of  $F$ . The argument above works for a sufficiently small neighborhood of the origin. However, we can simply relabel the origin and re-run the argument based at some other point in order to show that  $d_x F$  and  $d_y F$  exist in a small neighborhood of any point where  $F$  itself exists and satisfies the basic differential equation. This shows that  $d_x F$  and  $d_y F$  exist on exactly the same neighborhood that  $F$  exists on. Now we can repeat the argument. Lemma 3.1 says that  $d_x F$  is a  $W$ -extension of  $d_x f$ , where  $f$  and  $W$  are both smooth. Hence  $d_x F$  is continuous, and both  $d_{xx} F$  exist and  $d_{yx} F$  exist. Likewise  $d_{xy} F$  and  $d_{yy} F$  exist. And so on. This shows that  $F$  is smooth in the neighborhood on which it exists and satisfies the differential equation.

## 4 Proof of Lemma 3.1

This is a painful exercise in the chain rule. Using Equation 1 we have

$$\begin{aligned} \widehat{F}_h(x, y, s) &= \frac{F(x+h, y, s) - F(x, y, s)}{h} = \\ &= \frac{F(x+h, y, 0) - F(x, y, 0)}{h} + \\ &+ \int_0^s \frac{V(F(x+h, y, t)) - V(F(x, y, t))}{h} dt = \int_0^s R_1(s) ds. \end{aligned} \quad (2)$$

Now let's do something different.

Holding  $s$  fixed, let

$$\Psi(u) = V(u(F(x, y, s)) + (1-u)F(x+h, y, s)).$$

From the Fundamental Theorem of Calculus, we have

$$\begin{aligned} \Psi(1) - \Psi(0) &= \int_0^1 d_u(\Psi) du = \\ &= \int_0^1 d_x V(F(x+uh, y, s)) h \widehat{F}_{h,1}(x, y, s) + \\ &+ \int_0^1 d_y V(F(x+uh, y, s)) h \widehat{F}_{h,2}(x, y, s). \end{aligned} \quad (3)$$

This last line comes from the chain rule. we have set  $\widehat{F}_h = (\widehat{F}_{h,1}, \widehat{F}_{h,2})$ .

Note that  $\Psi(1) = V(F(x, y, t))$  and  $\Psi(0) = V(F((x + h, y, t))$ . Plugging this into the last equation and dividing by  $h$  we have

$$R(s) = \int_0^1 d_x V(F(x + uh, y, s)) \widehat{F}_{h,1}(x, y, s) du + \int_0^1 d_y V(F(x + uh, y, s)) \widehat{F}_{h,2}(x, y, s) du.$$

Now we define

$$A_{11}(F(x, y), s) = \int_0^1 d_x V(F(x + uh, y, s)) du,$$

and

$$A_{12}(F(x, y), s) = \int_0^1 d_y V(F(x + uh, y, s)) du.$$

With these definitions, we have

$$R_1(s) = A_{11}(F(x, y), s) \times \widehat{F}_{h,1}(x, y, s) + A_{12}(F(x, y), s) \times \widehat{F}_{h,1}(x, y, s). \quad (4)$$

Combining Equations 2 and 4 we see that

$$\widehat{F}_{h,1}(x, y, s) = \widehat{F}_{h,1}(x, y, 0) + \int_0^s A_{11}(F(x, y), t) \times \widehat{F}_{h,1}(x, y, t) + A_{12}(F(x, y), t) \times \widehat{F}_{h,1}(x, y, t) dt.$$

We have done all this with respect to the  $x$ -coordinate. Re-doing everything with respect to the  $y$ -coordinate, we have

$$\widehat{F}_{h,2}(x, y, s) = \widehat{F}_{h,1}(x, y, 0) + \int_0^s A_{21}(F(x, y), t) \times \widehat{F}_{h,1}(x, y, t) + A_{22}(F(x, y), t) \times \widehat{F}_{h,1}(x, y, t) dt.$$

Combining these last two equations into a matrix equation, we have

$$\widehat{F}_h(x, y, s) = \widehat{F}_h(x, y, 0) + \int_0^s A(F(x, y), t) \widehat{F}_h(x, y, t) dt.$$

This last equation is a matrix valued equation. This is an integral form of the basic differential equation, a special case of Equation 1. This completes the proof of Lemma 3.1.