Existence and uniqueness of solutions to an Initial Value Problem

Theorem 1. *The initial value problem $y' = H(y,t)$, $y(t_0) = y_0$ has a unique solution $y : [t_0 - \epsilon, t_0 + \epsilon] \to \mathbb{R}^k$ for small enough $\epsilon > 0$ if H is nice enough. To be more precise, "H nice enough" means that H is continuous and there is a neighborhood U of (y_0, t_0) and a constant M so that $||H(y,t) - H(z,t)|| \le M||y - z||$ for all (y,t) and (z,t) in U. (This is known as being Lipschitz in y.) For example, if H is differentiable with continuous derivatives it is nice enough.*

For example, the IVP $y' = |y|$, $y(0) = 0$ must have a unique solution since $H(y,t) = |y|$ is Lipschitz. But $y' = y^{1/3}$, $y(0) = 0$ has two solutions $y_1(t) = 0$ and $y_2(t) = \begin{cases} (2t/3)^{3/2} & \text{for } t \ge 0 \\ 0 & \text{for } t \le 0 \end{cases}$. The above theorem does not apply since $H(y,t) = y^{1/3}$ is not Lipschitz, essentially because its derivative approaches ∞ as $y \to 0$.

Proof: Let $X(\epsilon, \delta)$ be the space of continuous functions $f\colon [t_0 - \epsilon, t_0 + \epsilon] \to \mathbb{R}^k$ so that $||f(t) - y_0|| \le \delta$ for all $t \in [t_0 - \epsilon, t_0 + \epsilon]$. We have a map $\tau\colon X(\epsilon,\delta) \to X(\epsilon,\delta)$ given by $\tau(f)(t) = y_0 + \int_{t_0}^t H(f(s), s)\, ds$. For example, if $k = 1$, $t_0 = 0$, $y_0 = 1$, $f(t) = t^3$, and $H(y,t) = y^2 - t^2$ then $\tau(f)$ is the function g so that

$$g(t) = 1 + \int_0^t H(f,s)\, ds = 1 + \int_0^t s^6 - s^2\, ds = 1 + t^7/7 - t^3/3$$

The idea of the proof is first that solutions of the IVP are the same as fixed points of τ, that is, functions y so that $y = \tau(y)$. This is true since $d\tau(y)/dt = H(y,t)$. Next we show that τ has a unique fixed point. To do this, choose any function y_1 in $X(\epsilon,\delta)$, for example the constant function $y_1(t) = y_0$. Then let $y_2 = \tau(y_1)$, $y_3 = \tau(y_2)$, etc., so $y_{n+1} = \tau(y_n)$. If we can show the functions y_n approach a limit y then

$$\tau(y) = \lim_{n\to\infty} \tau(y_n) = \lim_{n\to\infty} y_{n+1} = y$$

So we have at least one solution y. We show it is unique by showing τ is contracting, that $||\tau(y) - \tau(z)|| < ||y - z||$ if $y \ne z$.

Now for the gory details. Pick $\delta > 0$ so $\{(x,t) \in \mathbb{R}^k \times \mathbb{R} \mid ||x - y_0|| \le \delta, |t - t_0| \le \delta\} \subset U$. Choose N so $||H(x,t)|| \le N$ for all (x,t) with $||x - y_0|| \le \delta$ and $|t - t_0| \le \delta$. Choose $\epsilon > 0$ so $\epsilon < \delta$, $\epsilon < 1/(2M)$ and $\epsilon < \delta/N$. If f is in $X(\epsilon,\delta)$ define $||f||$ to be the maximum of $||f(t)||$ for $t \in [t_0 - \epsilon, t_0 + \epsilon]$. Thus $||f - g||$ measures how far apart two functions are.

We have skipped over the well-definedness of τ. If $f \in X(\epsilon,\delta)$ is there any guarantee that $\tau(f) \in X(\epsilon,\delta)$ also? In particular, is $||\tau(f)(t) - y_0|| \le \delta$ for all t? We have

$$||\tau(f)(t) - y_0|| = ||\int_{t_0}^t H(f,s)\, ds|| \le |\int_{t_0}^t ||H(f,s)||\, ds| \le |\int_{t_0}^t N\, ds| = N|t - t_0| \le N\epsilon < \delta$$

so τ is well defined.

Pick $f, g \in X(\epsilon,\delta)$. There is a $t_1 \in [t_0 - \epsilon, t_0 + \epsilon]$ so $||\tau(f)(t_1) - \tau(g)(t_1)|| = ||\tau(f) - \tau(g)||$. We have

$$||\tau(f) - \tau(g)|| = ||\tau(f)(t_1) - \tau(g)(t_1)|| = ||\int_{t_0}^{t_1} H(f,s) - H(g,s)\, ds||$$

$$\le |\int_{t_0}^{t_1} ||H(f,s) - H(g,s)||\, ds| \le |\int_{t_0}^{t_1} M||f(s) - g(s)||\, ds|$$

$$\le |\int_{t_0}^{t_1} M||f - g||\, ds| = |t_1 - t_0|M||f - g|| \le M\epsilon||f - g||$$

So $||\tau(f) - \tau(g)|| \le ||f - g||/2$ since we have chosen ϵ so $M\epsilon < 1/2$.

Notice that

$$||y_{n+1} - y_n|| = ||\tau(y_n) - \tau(y_{n-1})|| \le ||y_n - y_{n-1}||/2$$

Thus for any $t \in [t_0 - \epsilon, t_0 + \epsilon]$ we have $||y_{n+1}(t) - y_n(t)|| \leq ||y_{n+1} - y_n|| \leq A/2^{n-1}$ where $A = ||y_2 - y_1||$. Consequently $\lim_{n\to\infty} y_n(t)$ exists since $y_n(t) = y_1 + \Sigma_{i=1}^{n-1} y_{i+1}(t) - y_i(t)$ and the series $\Sigma_{i=1}^{\infty} y_{i+1}(t) - y_i(t)$ converges absolutely by comparison with with $\Sigma_{i=1}^{\infty} A/2^{n-1} = 2A$. Thus we may define a limit of the functions y_n by $y(t) = \lim_{n\to\infty} y_n(t)$. Note that

$$||y(t) - y_n(t)|| = ||\Sigma_{i=n}^{\infty} y_{i+1}(t) - y_i(t)|| \leq \Sigma_{i=n}^{\infty} ||y_{i+1}(t) - y_i(t)|| \leq \Sigma_{i=n}^{\infty} A/2^{i-1} = 2^{2-n} A$$

Then for any n we get:

$$||(y(t+h) - y(t))/h - H(y(t),t)|| = ||(y(t+h) - y_n(t+h) + y_n(t+h) - y_n(t) + y_n(t) - y(t))/h - H(y(t),t)||$$

$$\leq ||y(t+h) - y_n(t+h)||/|h| + ||(y_n(t+h) - y_n(t))/h - H(y(t),t)|| + ||y_n(t) - y(t)||/|h|$$

$$\leq ||(y_n(t+h) - y_n(t))/h - H(y(t),t)|| + 2^{3-n}A/|h| = ||\int_t^{t+h} H(y_{n-1}(s),s) - H(y(t),t)\, ds||/|h| + 2^{3-n}A/|h|$$

$$\leq |\int_t^{t+h} ||H(y_{n-1}(s),s) - H(y(t),t)||\, ds/h| + 2^{3-n}A/|h|$$

Now

$$||H(y_{n-1}(s),s) - H(y(t),t)|| \leq ||H(y_{n-1}(s),s) - H(y_{n-1}(t),s)|| + ||H(y_{n-1}(t),s) - H(y(t),t)||$$

$$\leq M||y_{n-1}(s) - y_{n-1}(t)|| + ||H(y_{n-1}(t),s) - H(y(t),t)||$$

$$= M||\int_t^s H(y_{n-2}(u),u)\, du|| + ||H(y_{n-1}(t),s) - H(y(t),t)|| \leq MN|h| + ||H(y_{n-1}(t),s) - H(y(t),t)||$$

Pick any $\epsilon' > 0$. By continuity of H, if h is small enough and n is large enough, then

$$MN|h| + ||H(y_{n-1}(t),s) - H(y(t),t)|| < \epsilon'/2$$

for all s between t and $t+h$. But then as long as we choose n so $n > 4 + \log_2(A/(|h|\epsilon'))$ we are guaranteed that $||(y(t+h) - y(t))/h - H(y(t),t)|| < \epsilon'$. Consequently, $y(t)$ is differentiable and $y'(t) = \lim_{h\to 0}(y(t+h) - y(t))/h = H(y(t),t)$.

So we have at least one solution y. But if $z(t)$ is another solution to the IVP we have $||y - z|| = ||\tau(y) - \tau(z)|| \leq ||y - z||/2$ and thus $||y - z|| = 0$ so $y = z$. ∎

The only loose end is showing that if H is continuously differentiable then it is Lipschitz in y. But first a definition. If A is a matrix let $||A||$ be the maximum value of $||Ax||$ for all x with $||x|| = 1$. Then we know for any x that $||Ax|| \leq ||A||\,||x||$.

Choose a ball U centered at (y_0, t_0) and contained in the domain of H. Pick any (y,t) and (z,t) in U. Let $\alpha(s) = (sy + (1-s)z, t)$ be the path between these two points. We have

$$H(y,t) - H(z,t) = H(\alpha(1)) - H(\alpha(0)) = \int_0^1 dH(\alpha(t))/dt\, dt = \int_0^1 DH\,\alpha'(t)\, dt = \int_0^1 DH\,(y - z, 0)\, dt$$

Thus

$$||H(y,t) - H(z,t)|| \leq \int_0^1 ||DH\,(y - z, 0)||\, dt \leq \int_0^1 ||DH||\,||y - z||\, dt \leq M||y - z||$$

if M is the maximum of $||DH||$ on U.

Linear algebra note: In fact $||A||$ is the square root of the maximum eigenvalue of the symmetric matrix $A^T A$. To prove this: Choose an orthogonal matrix Q so $QA^TAQ^T = D$ a diagonal matrix. If $||x|| = 1$ then $||Ax||^2 = x^T A^T A x = x^T Q^T D Q x$. Setting $y = Qx$, letting d be the maximum of the diagonal entries d_i of D, we have

$$||Ax||^2 = y^T D y = \Sigma d_i y_i^2 \leq d \Sigma y_i^2 = d$$

since $||y|| = ||x|| = 1$. On the other hand, if v is the eigenvector for d of unit length then $||Av||^2 = v^T A^T A v = v^T d v = d v^T v = d$.