We want to show that if we have an equation of the form $f(x, y) = c$ (where f is differentiable with a continuous derivative as a function of x and y) then in an interval around any non-degenerate (we will see what this means) point, we have something where $\frac{dy}{dx}$ exists, ie the resulting curve is actually smooth.

Definition 1. A function is lipschitz continuous if we always satisfy that $|f(x) < f(y)| < k|x - y|$ for some k. This intuitively means its slope is never greater than k, however it need not be differentiable, it just can't ever be "vertical" or approach being vertical.

Definition 2. A family of functions F is Equicontinuous at a point x_0 if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all functions f in F,

$$|x_0 - x| < \delta \Rightarrow |f(x_0) - f(x)| < \varepsilon$$

Definition 3. Partial derivatives

We can have functions of multiple variables, like $f : \mathbb{R}^m \to \mathbb{R}^n$. As an example, suppose $z = x^2 + y^3$, then we can sketch this using a contour plot, kind of like elevation maps, where we show lines on the x,y-plane corresponding to when $x^2 + y^3$ is constant. Here is that example (Figure 1):

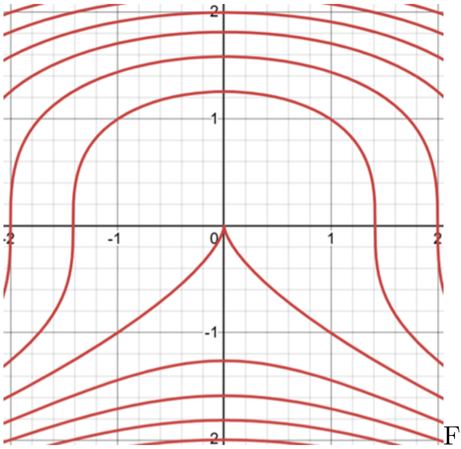

Figure 1

However, if I imagine this as a 3d graph of a surface with height equal to $x^2 + y^3$, then if I pick a point on this graph and try to find the slope, I have a problem that the slope depends on the direction. Therefore I write $\frac{\partial z}{\partial x}$ for the slope as I move in just the x direction and hold y constant. This is called a partial derivative. In this example, that is 2x, because we differentiate $x^2 + y^3$ and the y^3 vanishes since it is a constant. We put a little thingy in the corner like this to show what's being held constant, as shown below.

$$\left.\frac{\partial z}{\partial x}\right|_y$$

We do need to be careful about showing what is constant in some cases, as for example if f is a function of x, y and z, then $\left.\frac{\partial f}{\partial x}\right|_y$ does not always equal $\left.\frac{\partial f}{\partial x}\right|_z$. For example, in the surface $x^2+y^3+z^4 = 1$, then $\left.\frac{\partial f}{\partial x}\right|_y = \frac{d}{dx}\left(x^2 + y^3 + z^4\right) = 2x+4z^3\frac{dz}{dx}$ since y is constant, however $\left.\frac{\partial f}{\partial x}\right|_z = 2x + 3y^2\frac{dy}{dx}$.

Formally, for example, if z is a function of x and y, then $\left.\frac{\partial z}{\partial x}\right|_y$

Is defined as $\lim_{h\to 0} \frac{z(x+h,y)-z(x,y)}{h}$.

Example. $f(x, y) = x^2 + y^3 + e^{xy^2}$. As a shorthand for the partial derivative with respect to x we often instead write f_x. Since y is treated as constant here, we get that $f_x = 2x + y^2e^{xy^2}$, and $f_y = 3y^2 + 2xye^{xy^2}$.

Lemma 1. (Stone weierstrass theorem for real functions on 2D rectangles) For any continuous real function defined on a closed rectangle $[a, b]\times[c, d]$ we can define a sequence of Lipschitz continuous functions that converges uniformly to our function.

Proof. Fix $\varepsilon > 0$. We will use the known fact (Level 6.2) that our function is uniformly continous since it is continuous on a closed interval. So let δ be such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\varepsilon}{2}$ (we know this δ exists exactly by this level 6 result). Now what we will do is split our rectangle into rectangles with long side shorter than $\frac{\delta}{\sqrt{2}}$ (actually we just need that the longest diagonal is shorter than δ). Now in this rectangle we will define our function g to coincide with f at the corners and then change linearly between the corners. Now for any t in this rectangle with x_0 one of the rectangle's corners and y_0 its opposite corner, $|x_0 - t| < \delta \Rightarrow |f(x_0) - f(t)| < \varepsilon$, and also $|y_0 - t| < \delta \Rightarrow |f(y_0) - f(t)| < \varepsilon$. Therefore what we have is that $f(x_0)$ and $f(y_0)$ are within a band of width 2ε around $f(t)$, so thus $g(t)$ which is between $g(x_0)$ and $g(y_0)$ is in that band, so we must have $|g(t) - f(t)| < \varepsilon$. Therefore if we let $\varepsilon_n \to 0$, we will get uniform convergence, and all of these individual functions are lipschitz continuous because of how we defined them: There slope is bounded by $\sqrt{g_x^2 + g_y^2}$, where these slopes are finite because they are the finite change in the function across the rectangle divided by the non-zero width of the rectangle.

□

Lemma 2. (Arzela-Ascoli theorem specialized to real functions on 2D closed intervals) Let a sequence of functions f_n be uniformly bounded on a closed bounded interval (For our purposes, we will prove this for 1D or 2D intervals) and equicontinuous, then there exists a subsequence f_{n_k} that converges uniformly to a continuous function f.

Proof. Enumerate the rational points in $\mathbb{R}^2$. To do this, first enumerate the rationals, ie write them out in a list so that we have a one-to-one mapping between the positive integers and the rational numbers. There are many ways to do this, but one way is to go around like this image below where the x is the numerator and the y is the denominator, but exclude duplicates or 0 denominators.

Figure 2 shows a visual idea of how we will do this

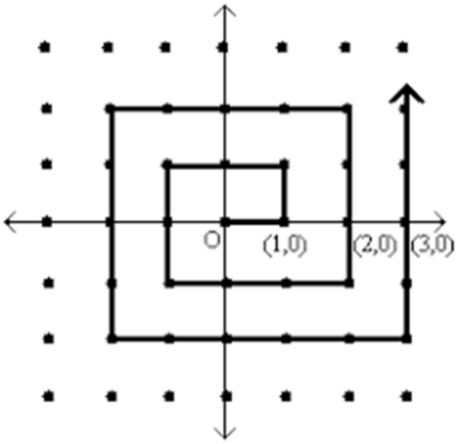

Figure 2

Then we know from level 6 technical results that the cartesian product of two listable sets is listable, in fact by a similar diagram to the image above.

Now we filter this enumeration so we only have the rational numbers in our interval. So we have a list of all the rational points in our interval. Call this enumeration $x_1, x_2, \ldots$

Since f_n has a uniform bound M, there is a sequence $f_{n_{1,k}}$ such that $f_{n_{1,k}}(x_1)$ converges pointwise by Bolanzo-Weierstrass (Level 4). We can find a further subsequence $f_{n_{2,k}}$ of this such that $f_{n_{2,k}}(x_2)$ also converges. We can get an infinite chain of subsequences this way. Now we want to form a sequence of functions f_k defined by $f_k = f_{n_{k,k}}$. By construction, this converges at every rational point. Therefore, given any $\varepsilon > 0$ and any rational point x_k, we can find an integer N such that for all n, m>N, we have that $|f_n(x_k) - f_m(x_k)| < \frac{\varepsilon}{3}$. We're making progress.

Since the family F is equicontinuous, there must be an open interval around x_k such that for any s and t in that open interval, $|f(s) - f(t)| < \frac{\varepsilon}{3}$ for all f in our family of functions. Doing this for all x_k gives a covering of our interval using open sets. We will prove shortly the fundamental result that this must admit a finite subcover since the interval is closed, but first I will remark that you can see that if this is true then the theorem about uniform continuity will follow since we can pick delta to be less than the smallest thing in this subcover.

Ok so suppose we have a closed bounded subset of $\mathbb{R}^n$ and suppose it is contained in a cube of side length M (possible by boundedness) and suppose we cover this closed bounded subset by open sets, then we want to show that we can get a finite subcover. So here is what we do: We take this cube and take our set S and split it into 2^n subcubes each of side length $\frac{M}{2}$, then one of these cubes intersected with S is a closed set where the part of our open cover that covers that set has no finite subcover – if all 2^n such intersections had a finite subcover so would the whole thing. Now take this subcube which we know exists and split it into subcubes again. When we iterate this, it is easy to see that we will converge to a single point in S (S is closed so it contains its limit points). Now this point contains an open set as part of the cover, but recall that open means every point has some small open ball centered at that point that is in the set. So if we take a subcube in the sequence that is much smaller than this open ball, it will be contained entirely within this open ball. So clearly, S intersected with this subcube has a finite subcover from the original open cover – in fact it is not only finite but covered by a single open set. This is a contradiction, so we have now proven the theorem.

Now with that important theorem done that we will use extensively in later levels, we will go back to our actual problem. There also exists an integer K such that each of these open sets in our subcover contains one of the first K rationals in our list, otherwise our list would be missing every rational in that interval. Finally, for any t in our interval, there are j and k<K such that t and x_k belong to the same interval U_j. For this choice of k, we have, by the triangle ineuqality:

$$|f_n(t) - f_m(t)| \leq |f_n(t) - f_n(x_k)| + |f_m(x_k) - f_n(x_k)| + |f_m(t) - f_m(x_k)|$$

Where we pick n and m to be at least as large as N which is at least large enough such that for all k from 1 to K we have the above inequality $|f_n(x_k) - f_m(x_k)| < \frac{\varepsilon}{3}$, and also from how we defined the U's, it is always the case that

$$|f_n(t) - f_n(x_k)|, |f_m(t) - f_m(x_k)| < \frac{\varepsilon}{3}$$

Therefore what we have is that for any t and fixed ε, $|f_n(t) - f_m(t)| \leq \varepsilon$ for m and n large enough. Therefore at each t, we see that the functions value is forced into an arbitrarily small band and thus we have pointwise convergence. We define f to be this pointwise limit, and then $|f_n(t) - f_m(t)| \leq \varepsilon$. Now letting m go to infinity and taking a pointwise limit, $|f_n(t) - f(t)| \leq \varepsilon$, so our sequence of functions converges uniformly. Since ε was arbitrary, we can say that, for example, $|f_n(t) - f(t)| \leq \frac{\varepsilon}{2} < \varepsilon$ so the inequality is strict. However, I need to show that a uniform limit of continuous functions is continuous.

Fix $\varepsilon > 0$. Pick N such that for all n>N, and all t $|f_n(t) - f(t)| < \frac{\varepsilon}{3}$. By continuity, for each x we have that

$$|x - y| < \delta \Rightarrow |f_N(x) - f_N(y)| < \frac{\varepsilon}{3}$$

for some δ. Then, by the triangle inequality,

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f(y) - f_N(y)| < \varepsilon.$$

□

Lemma 3. (Peano existence theorem) Let $f(x, y)$ be continuous on an open interval D around (x_0, y_0), then the differential equation $y'(x) = f(x, y)$ with initial condition (x_0, y_0) has a solution in a neighbourhood about that point that is not necessarily unique. (It is unique under another mild condition, in fact this is when f is Lipschitz continuous, but we do not need that so we will just do the existence theorem for now).

Proof. By replacing y with $y - y_0$ and similarly for x, we can assume that the initial condition is that we must pass through the origin. Since D is open, define a closed rectangle $R := [-x_1, x_1] \times [-y_1, y_1]$ contained in D. On R, the extreme value theorem implies that $\sup_R |f| \leq C < \infty$. Now by Lemma 1, pick a sequence of lipschitz continuous functions f_n converging uniformly to f with $\sup_R |f_k| \leq 2C < \infty$. We define the picard itrations $y_{k,n} : I = [-t_2, t_2] \to \mathbb{R}$ where we set that $t_2 = \min\left(t_1, \frac{y_1}{2C}\right)$, as follows:

$y_{k,0}(t) = 0$ and $y_{k,n+1}(x) = \int_0^x f_k(t, y_{k,n}(t))\, dt$. They are well defined by induction as we have that

$$|y_{k,n+1}(x)| \leq \int_0^x |f_k(t, y_{k,n}(t))|\, dt \leq |x| \sup_R |f_k| \leq t_2 2C \leq y_1$$

and thus $(t, y_{k,n}(t))$ is within the domain of f_k. Also, by the triangle inequality for integrals,

$$|y_{k,n+1}(x) - y_{k,n}(x)| = \int_0^x |f_k(t, y_{k,n}(t)) - f_k(t, y_{k,n-1}(t))|\, dt \leq L_k \int_0^x |(t, y_{k,n}(t)) - (t, y_{k,n-1}(t))|\, dt$$

Where for each k, an L_k exists by the Lipschitz condition.

Now define

$$M_{k,n}\left(x\right)=\sup_{t\in[0,x]}\left|y_{k,n+1}\left(t\right)-y_{k,n}\left(t\right)\right|\leq L_k\int_0^x M_{k,n-1}(t)dt$$

We also have that

$$M_{k,0}\left(x\right)=\sup_{t\in[0,x]}\left|y_{k,1}\left(t\right)-y_{k,0}\left(t\right)\right|=\sup_{t\in[0,x]}\left|y_{k,1}\left(t\right)\right|\leq\int_0^x\left|f_k\left(t,0\right)\right|dt\leq 2C|x|$$

Now we will prove by induction what we have the following bound for x in I for which we just proved the base case:

$$M_{k,n}\left(x\right)\leq\frac{\left(2CL_k^n|x|^{n+1}\right)}{(n+1)!}$$

Lets do the induction step. Suppose this is true, then

$$M_{k,n+1}\left(t\right)\leq L_k\int_0^x M_{k,n}\left(t\right)dt\leq L_k\int_0^x\frac{\left(2CL_k^n|x|^{n+1}\right)}{(n+1)!}dt\leq 2CL_k^{n+1}\int_0^x\frac{\left(|x|^{n+1}\right)}{(n+1)!}dt=\frac{\left(2CL_k^{n+1}|x|^{n+2}\right)}{(n+2)!}$$

As required. Crucially, this tends to 0 as n goes to infinity for all fixed x.

Now for x and x' in I, $\left|y_{k,n+1}\left(x'\right)-y_{k,n+1}\left(x\right)\right|\leq\int_x^{x'}\left|f_k\left(t,y_{k,n}\left(t\right)\right)\right|dt\leq 2C|t'-t|$, and thus since this always holds, the family of functions $y_{k,n}$ is equicontinuous: For ε given pick $\delta=\frac{\varepsilon}{2C}$. Therefore by lemma 2, for each k, there is a subsequence y_{k,a_n} converging uniformly to a continuous function y_k.

$$\left|y_{k,a_n}\left(x\right)-\int_0^x f_k\left(t,y_{k,a_n}\left(t\right)\right)dt\right|=\left|y_{k,a_n}\left(x\right)-y_{k,a_n+1}\left(x\right)\right|\leq M_{k,a_n}\left(x\right)\to 0$$

Thus, for each fixed x, we conclude $y_k\left(x\right)=\int_0^x f_k\left(t,y_k\left(t\right)\right)dt$ since the limit of $y_{k,a_n}\left(x\right)$ must coincide according to the inequality above, and the integral approaches $\int_0^x f_k\left(t,y_k\left(t\right)\right)dt$ since f is continuous so we can pass the y limit through it and then we are bounded by 2C so we can use dominated convergence (level 6 technical results) to pass the limit through the integral. Since each $y_{k,n}$ has a uniform bound y_1, so does y_k (the limit of those. Now, by the triangle inequality for integrals,

$$\left|y_k\left(x\right)-y_k\left(x'\right)\right|\leq\int_{x'}^x\left|f_k\left(t,y_k\left(t\right)\right)\right|dt\leq 2C|x-x'|$$

So y_k is equicontinuous so it has a subsequence k_n that converges uniformly to a continuous function y.

$y_{k_n}\left(x\right)=\int_0^x f_{k_n}\left(t,y_{k_n}\left(t\right)\right)dt$, therefore (supposing x is positive since the other way around is the same just with a sign flipped), $\int_0^x f_{k_n}\left(t,y_{k_n}\left(t\right)\right)dt\to\int_0^x f\left(t,y\left(t\right)\right)dt$ by the dominated convergence theorem and continuity of f. The limits on each side must coincide, so $y\left(x\right)=\int_0^x f(t,y)dt$. By the fundamental theorem of calculus, y is our solution to the differential equation so we are done at last.

□

Theorem. (Implicit function theorem in 2 dimensions) If f is continuously differentiable in a neighbourhood of a point (x_0,y_0) and $f_y(x,y)\neq 0$ in that neighbourhood, then there exists a unique differentiable function g such that

$$y_0=g\left(x_0\right),\ f\left(x,g\left(x\right)\right)=0$$

in a neighbourhood of x_0.

Proof. Lets try to find a g that works. By differentiating the equation $f\left(x,g\left(x\right)\right)=0$ we get

$f_x+g'\left(x\right)f_y=0$ so $g'\left(x\right)=-\frac{f_x}{f_y}$. Since f is continuously differentiable, f_x and f_y are continuous, and since $f_y\neq 0$, $\frac{f_x}{f_y}$ is continuous, so by Lemma 3 such a function g exists and it is continuous and differentiable.

Also, $g(x)$ actually satisfies this equation:

- Define $h(x) := f(x, g(x))$
- By the chain rule,
$$h'(x) = f_x(x, g(x)) + f_y(x, g(x))g'(x) = f_x - f_y\frac{f_x}{f_y} = 0$$
- So h is constant. At x_0, $h(x_0) = f(x_0, y_0) = 0$, so $h(x) = 0$ near x_0.

Note that for some fixed x, $f(x, y)$ as a function of y is increasing or decreasing (and thus injective) in a neighbourhood since f is continuously differentiable and $f(x_0, y)$ is not flat in y at our point. Therefore we have that if two solutions are the different, $f\left(x_0, g_1\left(x\right)\right) = f\left(x_0, g_2\left(x\right)\right)$ so by injectivity the two solutions are the same, thus we have uniqueness.

At a stationary point, at least one of the partial derivatives is non-zero, so we can use that one to construct a differentiable local curve as above.

$\square$