1 Recap

- Differentiability of vector fields.
- Chain rule.

2 Second derivatives

Let $f(x,y) = \frac{xy(x^2-y^2)}{x^2+y^2}$ when $(x,y) = (0,0)$ and $f(0,0) = 0$. f_x, f_y clearly exist away from $(0,0)$ and equal $\frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2}, -\frac{x(y^4+4x^2y^2-x^4)}{(x^2+y^2)^2}$ respectively. At $(0,0)$, $f_x = f_y = 0$ continue to exist. We aim to compute f_{xy}, f_{yx} at $(0,0)$. $f_{yx}(0,0) = \lim_{k\to 0} \frac{f_x(0,k)-f_x(0,0)}{k} = \lim_{k\to 0} \frac{-k^5}{k^5} = -1$. Likewise, $f_{xy}(0,0) = 1$. Thus they may not be equal in general!
Clairut's theorem: Assume that $f : S \subset \mathbb{R}^2 \to \mathbb{R}$ is scalar field, and $(a,b) \in S$ is an interior point. Suppose f_x, f_y, f_{xy}, f_{yx} exist in a neighbourhood of (a,b) and f_{xy}, f_{yx} are continuous at (a,b). Then $f_{xy}(a,b) = f_{y,x}(a,b)$. In particular, for C^2 functions, the mixed partials are equal.

Local extrema: Recall that in one-variable calculus, it makes sense to ask where a continuous function $f : [a,b] \to \mathbb{R}$ assumes its maximum and minimum possible values (global or absolute extrema). This makes sense because of the extreme value theorem. Now such a function can achieve its global extrema either at the end-points a and b or somewhere inside. If, in addition, f is differentiable on (a,b), then wherever it attains a local extremum (that is, a local max is a point $x_0 \in (a,b)$ such that $f(x) \leq f(x_0)$ for all x near x_0; likewise for a local min), $f'(x_0) = 0$. So to find global extrema, it suffices to look at the end-points and the local extrema.
One question: Given a local extremum, how can we decide whether it is a local max or a local min? To answer this question we need a better approximation (than the linear approximation that is).

Taylor theorem (second-order):
Recall that if f is differentiable at a then $f(x) = f(a) + f'(a)(x-a) + h_1(x)(x-a)$ where $h_1(x) \to 0$ as $x \to a$. If f is once-differentiable in $(a-\epsilon, a+\epsilon)$ for some $\epsilon > 0$, and twice-differentiable at a then Taylor's theorem holds: $f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + h_2(x)(x-a)^2$ where $h_2(x) \to 0$ as $x \to a$.
Proof: Define $h_2(x) = \frac{f(x)-f(a)-f'(a)(x-a)-\frac{f''(a)}{2}(x-a)^2}{(x-a)^2}$. At this point, one may use L'Hopital's rule (yes, there is a rigorous version; no I am not going to bore you with it) twice to see the result. (The proof is easier (using the fundamental theorem of calculus and integration by parts) if we assume that f''' exists and is continuous in $[a,x]$.)

Theorem: Suppose f attains a local extremum at an interior point a. Assume that f is once-differentiable on $(a-\epsilon, a+\epsilon)$ for some $\epsilon > 0$, and twice-differentiable at a. Then $f'(a) = 0$ and if $f''(a) > 0$, a is a point of local min, and if $f''(a) < 0$ it is a point of local max.
Proof: The fact that $f'(a) = 0$ was already proven. Nonetheless, if $f'(a) \neq 0$, then suppose $f'(a) > 0$ (the other case is similar). Then $f(x) - f(a) = f'(a)(x-a) + (x-a)h(x)$

where $h(x) \to 0$ as $x \to a$. Hence, for x close enough to a, $f'(a) + h(x) > 0$ and hence if $x < a$, $f(x) < f(a)$ whereas if $x > a$, $f(x) > f(a)$. Thus f is not a local extremum. Now if $f''(a) > 0$, then $f(x) - f(a) = (x-a)^2(\frac{f''(a)}{2} + h_2(x))$ where if x is close enough to a, then $\frac{f''(a)}{2} + h(x) > 0$ and hence $f(x) \geq f(a)$. Thus it is a local min. Likewise if $f''(a) < 0$.

Of course, one can wonder what happens when $f''(a) = 0$ (for instance, $f(x) = x^3$ and $a = 0$). In that case, it need not be a local extremum at all. If it is given to be a local extremum, then one needs to invoke a higher-order Taylor theorem (but in some cases, all the derivatives at the point can be zero and yet one can have a local extremum!).
Let $f(x) = x^3 - 3x$ on $[-2, \frac{1}{2}]$. Find all local extrema of f and decide whether they are local maxima or minima. Moreover, find the global extrema of f. $f'(x) = 3x^2 - 3 = 0$ when $x = \pm 1$. On the given domain, $x = -1$ is the only point where $f'(x) = 0$. (By the way, points where $f'(x) = 0$ are called critical points.) To find global extrema, compare $f(-2) = -2$, $f(\frac{1}{2}) = -\frac{11}{8}$, $f(-1) = -1 + 3 = 2$. So f attains a global max at $x = -1$ and a global min at $x = -2$. To find out whether f attains a local max or min at $x = -1$, $f''(x) = 6x = 6 \times -1 < 0$ and hence a local max.

3 Extrema in more than one variable

A scalar field $f : S \subset\to \mathbb{R}^n \to \mathbb{R}$ is said to have an absolute/global maximum at $\vec{a} \in S$ if $f(\vec{r}) \leq f(\vec{a})$ for all $\vec{r} \in S$ and likewise for an absolute/global minimum. The number $f(\vec{a})$ is called the absolute/global maximum *value* of f on S.
Just as in one-variable, there is an extreme value theorem: If $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is continuous, S is a closed subset and S is bounded, i.e., there is a finite-sized closed ball containing S, then f is bounded and assumes its maximum and minimum somewhere in S. There is no easy way to reduce it to one-variable. You would have to go through the proof again.
Given this theorem, it makes sense to ask how to calculate the global extrema. So we need local extrema. f is said to have a *local* maximum at an interior point $\vec{a} \in S$ if $f(\vec{r}) \leq f(\vec{a})$ for all $\vec{r}$ lying in an open ball around $\vec{a}$ that is completely contained in S. Likewise for a local minimum.

Just as in one-variable calculus, to find the global extrema of a differentiable function, we need to find all local extrema and compare them to what happens on the boundary. The boundary is not merely a finite collection of points! That is what makes this harder!
Theorem: Let f be differentiable at a local extremum $\vec{a}$. Then $\nabla f(\vec{a}) = \vec{0}$.
Proof: Let $\|\vec{v}\| = 1$. Let $g(t) = f(\vec{a} + t\vec{v})$ be defined for all $|t| < r$ for some small enough r. Then g is differentiable at 0 and attains a local extremum there. Thus $g'(0) = \langle \nabla f(\vec{a}), \vec{v} \rangle = 0$. Since this fact is true for all $\vec{v}$, $\nabla f(\vec{a}) = \vec{0}$. Points where the gradient vanishes are called critical points.
Caution: If f is *not* differentiable at a point, such a point deserves special consideration. For instance, $|x|$ assumes a local min at 0 and it isn't differentiable there. (Unlike us, some books call points of non-diff. critical points.)