1 Recap

1. Higher derivatives and Clairaut's theorem.
2. Taylor's theorem.
3. Second derivative test.
4. Positive-definiteness.

2 Bump functions

Unfortunately, even if the Taylor series converges, it need NOT be equal to the function itself! Let $E(t) = e^{-1/t}$ when $t > 0$ and 0 when $t \leq 0$. It turns out that $E(t)$ is C^∞ everywhere (an exercise/Lee's book).
Theorem: There exists a smooth function $\chi : \mathbb{R}^n \to \mathbb{R}$ that satisfies

1. $0 \leq \chi \leq 1$.
2. $\chi = 1$ on $[-1, 1] \times [-1, 1] \ldots [-1, 1]$.
3. $supp(\chi)$ is contained in $[-2, 2] \times [-2, 2] \ldots$.

We can find a function that satisfies similar properties but is instead radially symmetric. Such functions are called bump functions.
Proof: We first construct χ for $n = 1$. Let $\zeta(t) = \frac{E(t)}{E(t)+E(1-t)}$. Now $\zeta(t) = 0$ for $t \leq 0$ and $\zeta(t) = 1$ for $t \geq 1$. Let $\eta(t) = \zeta(2 + t)\zeta(2 - t)$. This does the job when $n = 1$. If we want a spherically symmetric bump function in $\mathbb{R}^m$, simply define it as $\chi(x) = \eta(r)$. If we want a cylindrically symmetric bump function in $\mathbb{R}^m$, simply define it as $\chi(x) = \eta(x^1)\eta(x^2)\eta(x^3)\ldots$. $\square$

3 The Einstein summation convention

It is painful to keep using Σ every time we want to sum over. Einstein invented a convenient notation for us. "Sum over indices that are repeated above and below. Column vectors have superscripts and row vectors have subscripts."
Examples:

1. If v is a real row vector, and w is a real column vector, then $\langle v^T, w\rangle = v_i w^i$.
2. If A is an $m \times n$ matrix, and v is an $n \times 1$ vector, then $(Av)^i = A^i_j v^j$.
3. $Tr(A) = A^i_i$.
4. $D(f \circ g)^i_j = [Df]^i_k [Dg]^k_j$.
5. $Tr(AB) = A^i_j B^j_i = B^j_i A^i_j = Tr(BA)$.

If V is a vector space, basis vectors are denoted with *subscripts* (like $e_1, e_2, \ldots$) and components w.r.t a basis are denoted with superscripts (like $v = v^i e_i$). For the dual space, the indices are flipped, i.e., $e_1^* = e^1, \ldots$ and $\omega = \omega_i e^i$.

4 The inverse function theorem and its cousins

Consider $x^2 + y^2 = 1$. Is y a differentiable function of x? Of course not! $y = \pm\sqrt{1-x^2}$. So y is not even a *function* of x! Also, both functions are not differentiable at $x = \pm 1$. The best we can say is "near" (a, b) on the circle, we can differentiably solve for either y in terms of x or x in terms of y. We can also generalise this observation to $x^2 + y^2 = r^2$. Near say $(\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$, $y = \sqrt{r^2 - x^2}$ is a smooth function of (r, x). What about $x^{\sin(y)} + y^2 + e^{xy} = 1$? It is not clear whether we can solve for either variable in terms of the other. Even if we can, it is unlikely that we can write a simple formula for the solution.

More generally, given a C^k function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, we can ask, "Can we solve for y as a C^1 function of x, c from $F(x, y) = c$?" Since C^1 functions can be approximated by linear functions, let us at least try our luck when $F(x, y) = Ax + By$ where A, B are constant matrices (of what type?). So if $Ax + By = c$, can we solve for y in terms of x, c uniquely? $By = c - Ax$. Therefore, this question has an affirmative answer iff B is an invertible matrix.

Often in maths, general cases can be reduced to the simplest non-trivial special cases. Suppose $n = 0$, i.e., if $F(y) = c$, can we locally solve for y as a C^1 function of c? That is, when do local inverses exist and are differentiable, i.e., when is F a *local* diffeomorphism? Also, what is the derivative of F^{-1}? When $n = 1$, by monotonicity, this problem is not hard. In general, we need a theorem. We *expect* that it is good enough for DF_a to be invertible. Indeed, this expectation is true.

Theorem 1 (The inverse function theorem (IFT))**.** *Let $U, V \subset \mathbb{R}^n$ be open sets. Suppose $F : U \to V$ is a C^k function (where $k \geq 1$ can be ∞). If DF_a is invertible at $a \in U$, then there exist connected neighbourhoods $a \in U_a \subset U$, $F(a) \in V_a \subset V$ such that $F : U_a \to V_a$ is a C^k-diffeomorphism. Moreover, $DF^{-1}_{f(a)} = (DF_a)^{-1}$.*

You might have seen the case when $k = 1$. Given that case, the general case is an easy inductive application of the chain rule to the formula for the derivative of F^{-1}. The proof of this theorem can be done using Newton's iteration. (Optional fun fact: The proof shows that the same kind of a result holds when U, V are open subsets of a Banach space. While we won't need this fact, it is useful for PDE.)
Corollary: Suppose $U \subset \mathbb{R}^n$ is open and $F : U \to \mathbb{R}^n$ is C^k ($1 \leq k \leq \infty$). Assume that $\det(DF_a) \neq 0 \ \forall \ a \in U$. Then F is an open map, and if F is $1-1$, $F : U \to F(U)$ is a C^k-diffeomorphism.
Proof: Open map: The IFT implies that F is a local diffeo. Thus, if $W \subset U$ is an open set, then $W = \cup_a U_a$. Since $F(U_a)$ is open and $F(W) = \cup_a F(U_a)$, we see that $F(W)$ is open.
Diffeomorphism: F is invertible. Thus, every local inverse coincide with F^{-1} and hence F^{-1} is a C^k-diffeo. $\square$
Therefore, $x = r\cos(\theta), y = r\sin(\theta)$ is a diffeo on $(0, \infty) \times (0, 2\pi)$ (more generally, on any open subset away from $r = 0$ where it is $1-1$).