We review briefly some definitions and results from multivariable calculus.

DEFINITION. If $f : \mathbf{R}^n \to \mathbf{R}$ is a differentiable function, then its *gradient* is defined by

$$\nabla f(\mathbf{x}) = \frac{\partial f}{\partial x^1}\mathbf{e}_1 + \frac{\partial f}{\partial x^2}\mathbf{e}_2 + \cdots + \frac{\partial f}{\partial x^n}\mathbf{e}_n,$$

where $x^1, \ldots, x^n$ are the coordinates of the point $\mathbf{x}$ and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis for $\mathbf{R}^n$. In particular, if $n = 2$ we have

$$\nabla f(x, y) = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}.$$

If $\mathbf{F}$ is a vector field on $\mathbf{R}^n$, then we define its *divergence* to be the function

$$\operatorname{div}\mathbf{F}(\mathbf{x}) = \frac{\partial \mathbf{F}^1}{\partial x^1} + \cdots + \frac{\partial \mathbf{F}^n}{\partial x^n},$$

where $\mathbf{F}^1, \ldots, \mathbf{F}^n$ are the components of the vector field $\mathbf{F}$ in the standard basis on $\mathbf{R}^n$.

If $f : \mathbf{R}^n \to \mathbf{R}$ is a twice-differentiable function, then we define its *Laplacian* (named after the French mathematician Pierre Simon de Laplace) to be the function

$$\nabla^2 f(\mathbf{x}) = \operatorname{div}\nabla f(\mathbf{x}) = \frac{\partial^2 f}{\partial {x^1}^2} + \cdots + \frac{\partial^2 f}{\partial {x^n}^2}.$$

($\nabla^2 f$ is also sometimes written – particularly by mathematicians – as Δf.) (As an aside, we recall the definition of the curl of a vector field on $\mathbf{R}^3$:

$$\operatorname{curl}\mathbf{F} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\mathbf{i} - \left(\frac{\partial F_z}{\partial x} - \frac{\partial F_x}{\partial z}\right)\mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\mathbf{k}.$$

This will not be as important in this course as the other two operators above. Also we note that, unlike the two operators above, in dimensions higher than 3 the curl can no longer be defined as a vector field; it becomes instead what is known as a differential 2-form. The theory of differential forms is a very important part of modern mathematics, but is rather beyond the scope of the present course.)

DEFINITION. If $\mathbf{x} = (x, y) \in \mathbf{R}^2 \backslash \{(0, 0)\}$, we define its *polar coordinates* (r, θ) to be the real numbers $r \in (0, +\infty)$, $\theta \in [0, 2\pi)$ such that

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta. \end{aligned}$$

If $\mathbf{x} = (x, y, z) \in \mathbf{R}^3 \backslash \{(0, 0, z) | z \in \mathbf{R}\}$, then we (following Pinsky, see [1], p. 171) define its *cylindrical coordinates* (ρ, ϕ, z) to be the real numbers $\rho \in (0, +\infty)$, $\phi \in [0, 2\pi)$, $z \in \mathbf{R}$ such that

$$\begin{aligned} x &= \rho\cos\phi \\ y &= \rho\sin\phi \\ z &= z, \end{aligned}$$

and its *spherical polar coordinates* (also *spherical coordinates*, or *polar coordinates*) (r, θ, ϕ) to be the real numbers $r \in (0, +\infty)$, $\theta \in (0, \pi)$, $\phi \in [0, 2\pi)$ such that

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta. \end{aligned}$$

Note that the role played by θ in (planar) polar coordinates is played by ϕ in cylindrical and spherical coordinates. It is called the azimuthal angle, while the angle θ in spherical coordinates is called the polar angle (and is, for us, generally the more interesting of the two, when we are working with spherical coordinates, that is).

Definition. Let $\mathbf{F} : \mathbf{R}^n \to \mathbf{R}^n$ be differentiable at some point $\mathbf{x} \in \mathbf{R}^n$ (by which we mean that each of its components $\mathbf{F}^1, \ldots, \mathbf{F}^n$ is differentiable at that point). Then its *Jacobian* (named after the German mathematician Karl Gustav Jacob Jacobi) is defined to be the determinant

$$J(\mathbf{F})(\mathbf{x}) = \begin{vmatrix} \frac{\partial \mathbf{F}^1}{\partial x_1} & \frac{\partial \mathbf{F}^1}{\partial x_2} & \cdots & \frac{\partial \mathbf{F}^1}{\partial x_2} \\ \frac{\partial \mathbf{F}^2}{\partial x_1} & \frac{\partial \mathbf{F}^2}{\partial x_2} & \cdots & \frac{\partial \mathbf{F}^2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{F}^n}{\partial x_1} & \frac{\partial \mathbf{F}^n}{\partial x_2} & \cdots & \frac{\partial \mathbf{F}^n}{\partial x_n} \end{vmatrix},$$

where all partial derivatives are to be evaluated at the point $\mathbf{x}$.

Theorem. Let $\mathbf{F} : U \to \mathbf{R}^n$ be differentiable and one-to-one, where $U \subset \mathbf{R}^n$ is a product of intervals[1], and suppose that $J(\mathbf{F}) \neq 0$ on U except on a set of dimension lower than n[2]. Let $f : \mathbf{F}(U) \to \mathbf{R}$ be an integrable function. Then

$$\int \cdots \int_{\mathbf{F}(U)} f(\mathbf{x}) dx^1 \cdots dx^n = \int \cdots \int_U f(\mathbf{F}(\mathbf{y})) |J(\mathbf{F})(\mathbf{y})| dy^1 \cdots dy^n.$$

(Here $dx^1 \cdots dx^n$, etc., denotes the element of n-dimensional hypervolume; i.e., what we usually write as dA in $\mathbf{R}^2$, and dV in $\mathbf{R}^3$.)

When applied to polar, cylindrical, and spherical coordinates, this theorem gives the following formulæ. For simplicity, we shall use V to denote the region in $\mathbf{R}^2$ or $\mathbf{R}^3$ and U to denote the corresponding region in coordinate space. We shall also use the expressions $dxdy$, etc., to denote the relevant area or volume element for each integral.

$$\begin{aligned} \int_V f(x,y) dxdy &= \int_U f(r,\theta) r dr d\theta \\ \int_V f(x,y,z) dxdydz &= \int_U f(\rho,\phi,z) \rho d\rho d\theta dz \\ \int_V f(x,y,z) dxdydz &= \int_U f(r,\theta,\phi) r^2 \sin\theta dr d\theta d\phi \end{aligned}$$

.

The first homework assignment has some problems which make use of these formulæ.

Definition. Let $f : \mathbf{R}^n \to \mathbf{R}$. A point $\mathbf{x} \in \mathbf{R}^n$ is called a local maximum (respectively minimum) of f if $f(\mathbf{x}) \geq f(\mathbf{y})$ (respectively $f(\mathbf{x}) \leq f(\mathbf{y})$) for all $\mathbf{y}$ near[3] $\mathbf{x}$.

Let $f : \mathbf{R}^n \to \mathbf{R}$ have continuous second-order partial derivatives. Then the *Hessian* matrix of f at a point $\mathbf{x} \in \mathbf{R}^n$ is defined to be the matrix

$$H(f)(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^1 \partial x^1} & \frac{\partial^2 f}{\partial x^1 \partial x^2} & \cdots & \frac{\partial^2 f}{\partial x^1 \partial x^n} \\ \frac{\partial^2 f}{\partial x^2 \partial x^1} & \frac{\partial^2 f}{\partial x^2 \partial x^2} & \cdots & \frac{\partial^2 f}{\partial x^2 \partial x^n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x^n \partial x^1} & \frac{\partial^2 f}{\partial x^n \partial x^2} & \cdots & \frac{\partial^2 f}{\partial x^n \partial x^n} \end{pmatrix}.$$

Theorem. Let $f : \mathbf{R}^n \to \mathbf{R}$ be a differentiable function, and suppose that $\mathbf{x}$ is a local maximum or local minimum of f. Then $\nabla f(\mathbf{x}) = 0$.

---

[1] In general, it is sufficient for U to be what is called an *open set*; for those who haven't seen the concept of open sets, you can think of U as being a region bounded by suitably nice hypersurfaces.

[2] More specifically, except on a set of *content zero*; but we shall have no need to do things in such generality.

[3] I.e., in a neighborhood of.

THEOREM. Let $f : \mathbf{R}^n \to \mathbf{R}$ be a twice-differentiable function with continuous second-order partial derivatives, and suppose that $\mathbf{x} \in \mathbf{R}^n$ is such that $\nabla f(\mathbf{x}) = 0$. Then:

(a) If all of the eigenvalues of $H(f)(\mathbf{x})$ are positive, then f has a local minimum at $\mathbf{x}$.

(b) If all of the eigenvalues of $H(f)(\mathbf{x})$ are negative, then f has a local maximum at $\mathbf{x}$.

(c) If the eigenvalues of $H(f)(\mathbf{x})$ are of mixed sign, then f has a saddle point at $\mathbf{x}$.

(d) Otherwise (in other words, some of the eigenvalues of $H(f)(\mathbf{x})$ are zero), $\mathbf{x}$ can be a local maximum, local minimum, or saddle point of f.

Possibly you have never seen this result at quite this level of generality. The two-dimensional version is usually stated as a test in terms of two quantities, say $\frac{\partial^2 f}{\partial x^2}$ and $\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial f}{\partial x \partial y}\right)^2$. It is easy to see that this last quantity is just the determinant of the Hessian matrix; thus in the two-dimensional case it is positive exactly when the eigenvalues of $H(f)(\mathbf{x})$ are all of the same sign, i.e., when we are in case (a) or (b) above. A little thought allows us to reduce the above theorem to the typical test in this case, but it is a bit tedious to write out.

REFERENCES.

1. Pinsky, Mark A. Partial Differential Equations and Boundary-Value Problems with Applications, 3d ed. Providence, Rhode Island: American Mathematical Society, 2011.