0 Introduction to Multi-Variable Calculus

0.1 Multi-Variable Functions; Examples; Notation

Note that you have encountered your first example of a multi-variable function already in primary school, long before you knew about the concept of a function: that function's name was 'plus': E.g., $\text{plus}(3,5) = 8$. Of course they didn't *use* the function notation, but wrote $3+5$ instead of $\text{plus}(3,5)$, but it still is a function nevertheless. This goes to stress that there is really nothing eery about multi-variable functions, and the curricular separation between single variable and multi-variable calculus is artificial and, in some contexts, obfuscating things.

Physics Examples: Many useful functions are multi-variable; consider for instance the temperature as a function of four variables: latitude, longitude, altitude, time. Often, these variables represent space, or space and time. But not always: For instance you could consider the volume of a certain amount of gas as a function of pressure and temperature.

Notation: Since calculus arose at the interface of mathematics and physics, different cultures of notation have influenced it, and it is helpful to point these out explicitly to avoid confusion. We'll use the mathematical notation for functions, which has a distinctive syntactic flavor: a certain symbol like f denotes the name of the function. A function takes a list of arguments (say longitude λ, latitude φ, altitude h, and time t) as inputs, and its result may be a certain real number (representing the temperature T at a certain place and time). This number will be denoted $f(\lambda, \varphi, h, t)$. So we say $T = f(\lambda, \varphi, h, t)$ and distinguish the *value* of the function from the function itself. f is the function, $f(\lambda, \varphi, h, t)$ its value.

Physicists often prefer a different notation. In saying "temperature is a function of longitude, latitude, etc", they prefer to call temperature itself a function, and so they assign the symbol T for temperature both to a quantity (which is a real number) and to a function (which is a machine that devours a list of input numbers and returns a certain value as output). They might say T is the temperature (value of a function), and write $T(\lambda, \varphi, h, t)$ for the function. So their usage of the parentheses is the opposite of the mathematicians' usage! Physics usage is semantics focused. The same symbol may be used for different functions, because both functions produce the same physical quantity as output, but dependent on different variables: $V(p, T)$ gas volume depending on pressure and temperature, versus $V(E, T)$ gas volume depending on energy and temperature of the gas. The mathematician's notation would be $V = f(p, T)$ vs. $V = g(E, T)$ instead.

This observation may seem pedantic and boring, but getting it clear will help avoid confusion later (eg., when dealing with the MV chain rule and partial derivatives).

The plane $\mathbb{R}^2$, and brethren; graphing functions: Like single real numbers x can be geometrically represented as points on the real line, pairs of real numbers (x, y) can be represented as points in a plane, and triplets of real numbers (x, y, z) as points in space. If we want a graph of a 2-variable function, f, given by the equation $z = f(x, y)$, this graph will be a surface, such that above each point (x, y) in the domain of f, there lies exactly one point of the surface.

This geometric representation is not feasible for functions of 3 or more variables, but the two-variable case will provide a useful metaphor for the more general case. All key concepts generalize easily from the 2-variable case to the n-variable case, even if we lose the geometric representation of the graph.

Another way to represent functions f graphically is to draw the level sets. For two variables, these are curves $f(x,y) = z_i$ for a list of sample values z_i, and tagging the value z_i at the curves. This is how, e.g., the temperature function is represented on the weather maps in USA Today, or the height function on hiking maps.

0.2 Some useful coordinate systems

Polar coordinates: Rather than representing a point P in the plane in terms of cartesian coordinates (x,y), you can also describe it in terms of its distance r from the origin O, and the angle of the line OP to the x axis. Let's call this angle φ, measured counterclockwise from the x-axis. The connection between cartesian and polar coordinates is given by $x = r\cos\varphi$, $y = r\sin\varphi$.

Cylindrical coordinates in space: They are obtained by replacing the cartesian coordinates x, y with polar coordinates, but retaining the z coordinate. They are useful to describe situations that have a cylinder symmetry.

Spherical coordinates (geography variant): These coordinates consist of geographical longitude λ (which plays the role of φ in polar coordinates), geographical latitude ϕ, and distance from the origin R (which I write with a capital R to avoid confusion with the r from cylinder coordinates, which gives the distance from the z axis. The connection of spherical and cartesian coordinates is

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

Spherical coordinates (physics variant): Here we replace the geographical latitude ϕ with the (angular) distance from the 'North pole' $\vartheta = \frac{\pi}{2} = \phi$ and get

$$\begin{aligned} x &= R\sin\vartheta\,\cos\lambda \\ y &= R\sin\vartheta\,\sin\lambda \\ z &= R\cos\vartheta \end{aligned}$$

(Conventional symbols for these coordinates vary. Many use φ for λ, some swap the symbols ϑ and φ.) Spherical coordinates are convenient in situations with spherical symmetry.

0.3 Vectors

Vectors in the plane, geometric point of view: Think of a vector in the plane as an arrow connecting two points. The word stems from the Latin vehere, which means to drive or to carry. Vectors, by definition, can be moved around without turning. So neither the initial point alone, nor the end point alone, is part of the information prescribing a vector, but it is rather the relative location of the two with respect to each other that makes the vector.

Vectors in the plane, algebraic point of view: If a vector, let's call it $\vec{a}$, goes from point A with coordinates (x_1, y_1) to point B with coordinates (x_2, y_2), then it travels $\Delta x = x_2 - x_1$ in x direction and $\Delta y = y_2 - y_1$ in y direction. We can view the numbers Δx and Δy as coordinates that determine the vector. The convention is that we write vectors NOT in the form $(\Delta x, \Delta y)$, but rather as 'columns' $\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$.

The 'vertical' convention for vectors: This convention is sometimes disobeyed with impunity in calculus textbooks, but I will strictly adhere to it in this class. For the moment this convention serves the trivial purpose of distinguishing the geometrically different objects 'point' and 'vector' by a simple distinction in notation. The deeper significance of this convention will become clear only later, when we introduce matrices, multiplication of matrices, and of matrices with vectors. In the context of calculus, this happens when we discuss derivatives of multi-variable functions.

This will be the time where (popular) neglect of the stated conventions comes at the expense of blurring the geometric significance of derivatives. In advanced calculus, you *must* abide by the 'vertical' notation for vectors, whereas elementary calculus textbooks, the authors can get by with keeping things 'simple' by not bothering.

Truly, writing vector components above each other is a *typographical* nuisance. There is an easy cop-out which combines the typographical convenience of writing coordinates side by side with the benefits that the vertical convention is to reap later. $[\Delta x\,,\ \Delta y]^T$ is a synonym for $\left[\begin{smallmatrix}\Delta x\\ \Delta y\end{smallmatrix}\right]$. The ^{T} does *not* refer to any exponent, but reads as 'transpose' and simply means: change the horizontal into vertical. The transpose notation is also borrowed from matrices and will show up more naturally later, when we discuss these.

Adding vectors: Geometrically, vectors (arrows) $\vec{a}$ and $\vec{b}$ are 'added' by moving the tail of vector $\vec{b}$ at the tip of vector $\vec{a}$, and then connecting the tail of $\vec{a}$ to the tip of $\vec{b}$. This connecting arrow is now the vector $\vec{a}+\vec{b}$. Tongue-in-cheek, I'll call this the 'dog-sniff-dog' rule of vector addition, but this is not an 'official' term.

Algebraically, the vector $\vec{a}$ may be represented as $\left[\begin{smallmatrix}\Delta x_a\\ \Delta y_a\end{smallmatrix}\right]$ and $\vec{b} = \left[\begin{smallmatrix}\Delta x_b\\ \Delta y_b\end{smallmatrix}\right]$. Then $\vec{a}+\vec{b} = \left[\begin{smallmatrix}\Delta x_a+\Delta x_b\\ \Delta y_a+\Delta y_b\end{smallmatrix}\right]$. So vectors my be added by adding corresponding components.

It is easy to see geometrically, and also easy to see in terms of the algebraic representation, that the addition of vectors abides by similar rules as the addition of numbers: $\vec{a}+\vec{b} = \vec{b}+\vec{a}$, and $(\vec{a}+\vec{b})+\vec{c} = \vec{a}+(\vec{b}+\vec{c})$.

We also introduce the null vector $\vec{0} = \left[\begin{smallmatrix}0\\ 0\end{smallmatrix}\right]$, even though it cannot be drawn like a nice arrow geometrically (Tail and tip would have to be the same point, and there is no direction then).

Note that the correspondence between geometric and algebraic vectors depends on the choice of coordinate directions. If your neighboor chooses to call a different direction as the x direction, this will result in different numbers Δx and Δy being assigned to the (geometrically) same vector. So the identification between geometric and algebraic vectors depends on the choice of coordinates, but any choice of (cartesian) coordinates is equally good; only you need to keep this choice consistent through all your calculations.

Vectors in physics: In physics, vectors are 'quantities that have not only a magnitude but also a direction'. Velocity, and force are the most popular examples. The magnitude of the quantity (which in the case of velocity is known under the name of 'speed') corresponds to the length of the arrow, and the direction carries over immediately. There is something important hidden behind this identification between physical vectors and mathematical vectors. If you toss out a ball horizontally from a window and at the same time let an identical ball drop from the window (without giving it a thrust either up or down), then both balls will be at the same height at the same time. In other words, the horizontal part of the velocity does not interfere with the vertical part. This is an experimental fact of physics. If it were not true, then the mathematical concept of a vector, as outlined above, would not be useful to describe physical velocities, forces, etc.

Vectors versus points: If you choose a distinct point in the plane, call it O, and make it the origin of the coordinate system (so $O = (0,0)$), then any point P with coordinates (x,y) in the plane can be represented by a vector that goes from O to P. This vector is obviously $\begin{bmatrix} x-0 \\ y-0 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$. In calculus, we often blur the distinction between the point $P = (x, y)$ and the vector $\vec{OP} = \begin{bmatrix} x \\ y \end{bmatrix}$. In situations, where our variables never had a geometric meaning originally (e.g., they represent the pressure and the temperature of some gas) and where we have just made these quantitites into coordinates of some 'artificial' point for the sake of graphic representation, there is no harm in this non-distinction. One is as good (rather: as artificial) as the other. Physicists will (justly and indignantly) protest if you suggest to make pressure and temperature into components of a vector. If this ever happens implicitly in a multivariable calculus context despite such protest, it is due to an artificial overuse of the mathematical vector concept within this limited context, where such overuse does neither good nor harm.

In situations, where our variables had a geometric meaning originally, this geometric meaning gets obfuscated, when you choose not to distinguish (x, y) and $\begin{bmatrix} x \\ y \end{bmatrix}$. It is only by arbitrarily singling out one point as origin that an identification between points and vectors can be made.

For dot product and norm, refer to any calculus 3 textbook, or the parts of the linear algebra glossary I am handing out.

0.4 Limits and Continuity

Limits: Remember the definition of a limit from single variable calculus: $\lim_{x\to x_*} f(x) = L$ means: for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for $0 < |x - x_*| < \delta$ (and x in the domain of f) it holds $|f(x) - L| \leq \varepsilon$. — To avoid trivialities, we also require that every set $0 < |x - x_*| < \delta$ contains numbers x in the domain of f. In informal language, the definiton means that we can get $f(x)$ arbitrarily close to L is only we assure that x is sufficiently close by x_* (but different from x_*).

The absolute value of the difference of two numbers has the geometric meaning of a distance between these numbers (represented by points on the real line. With this understanding, the definition of limit carries over to the multi-variable setting. The distance between points P and Q in the plane (represented by vectors $\vec{p}$ and $\vec{q}$) is $\|\vec{p} - \vec{q}\| = \|\vec{QP}\|$.

Definition: Assume that every set $0 < \|\vec{PP_*}\| < \delta$ contains points P in the domain of f. Then, $\lim_{P\to P_*} f(P) = L$ means: for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for $0 < \|\vec{PP_*}\| < \delta$ (and P in the domain of f) it holds $|f(P) - L| \leq \varepsilon$.

Continuity: In the definition of a limit, f need not be defined for P_*. But if $f(P_*)$ is defined, then we can ask whether $f(P_*)$ is equal to the limit $\lim_{P\to P_*} f(P)$.

Definition: We say f is continuous at P_* if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for $\|\vec{PP_*}\| < \delta$ (and P in the domain of f), it holds $|f(P) - f(P_*)| \leq \varepsilon$.

If every set $0 < \|\vec{PP_*}\| < \delta$ intersects the domain of f, this is equivalent to saying $\lim_{P\to P_*} f(P) = f(P_*)$.

All the usual rules for limits carry over: $\lim_{P\to P_*}(f(P)+g(P)) = \lim_{P\to P_*} f(P)+\lim_{P\to P_*} g(P)$ provided both limits on the right side exist. The same statement applies for $-$, $\cdot$, $/$ instead of $+$, subject to the extra hypotheses $\lim_{P\to P_*} g(P) \neq 0$ in the case of division. Therefore, sums, differences, products and quotients of continuous functions at P_* are again continuous at P_*, subject to the stipulation that denominators must not vanish at P_*.

If the functions g_i are continuous at P_* and f is continuous at $Q_* = (g_1(P_*), \dots, g_n(P*))$, then the composite function h given by $h(P) = f(g_1(P), \dots g_n(P))$ is continuous at P_*.

The squeeze theorem carries over from single variable calculus verbatim.

Further generalization: So far, we have considered scalar valued functions of several variables: scalar valued means that the values of these functions are real numbers. We will also consider vector valued functions of either one or several variables; i.e., the values of these functions will themselves be vectors. This is very natural: For instance for a particle moving around in space, we can consider the position vector (i.e. the vector from the origin to the present location of the particle) as a function of time. Or we could consider the velocity of this particle as a function of time.

For instance, $\vec{p}(t) = \begin{bmatrix} \cos t \\ \sin t \\ 3t \end{bmatrix}$ would describe a particle moving upward as if along a spiral staircase. The time derivative of this function (of a single variable t) would be calculated componentwise: $\vec{p}'(t) = \begin{bmatrix} -\sin t \\ \cos t \\ 3 \end{bmatrix}$ and would describe the velocity of the particle.

Vector valued functions of several variables are a bit less intuitive, but consider the situation of water moving around in a container. Then the velocity $\vec{v}$ is a function of position (3 space coordinates (x, y, z)) and time t.

The same definitions for limit and continuity as before will again apply: the only distinction is that the limit L would also need to be a vector $\vec{L}$, and instead of $|f(P) - L|$, we'd have to write $\|\vec{f}(P) - \vec{L}\|$.

It is easy to see that $\lim_{P \to P_*} \vec{f}(P) = \vec{L}$ if and only if $\lim_{P \to P_*} f_i(P) = L_i$, where $f_i(P)$ and L_i are the coordinate components of the vectors $\vec{f}(P)$ and $\vec{L}$ respectively. And this applies, regardless of whether P stands for a single variable, or for several variables.

Examples:

(1) $\lim_{(x,y) \to (0,0)} \frac{x^2 y}{x^2 + y^2} = 0$ by the squeeze theorem, because $|\frac{x^2 y}{x^2 + y^2}| \le |y|$ and $\lim_{(x,y) \to (0,0)} |y| = 0$.

(2) In contrast, for $f(x, y) := \frac{xy}{x^2 + y^2}$, $\lim_{(x,y) \to (0,0)} f(x, y)$ does not exist. In every punctured disc $0 < \| \begin{bmatrix} x \\ y \end{bmatrix} \| < \delta$, there are points where $f(x, y) = \frac{1}{2}$ (e.g., $(x, y) = (\frac{\delta}{2}, \frac{\delta}{2})$), but also points where $f(x, y) = 0$ (e.g., $(x, y) = (\frac{\delta}{2}, 0)$). But the values of f cannot get close to 0 and to $\frac{1}{2}$ at the same time. Specifically, for $\varepsilon = \frac{1}{5}$, there can be no limit L such that $|L - \frac{1}{2}| \le \varepsilon$ and $|L - 0| \le \varepsilon$ at the same time.

Warning: If the 2-variable function f is continuous at every (x, y), then the one-variable functions g, given by $g(x) := f(x, y)$ for each fixed y, and h, given by $h(y) := f(x, y)$ for each fixed x are continuous. The converse however is NOT true. If $f(x, y) := \frac{xy}{x^2 + y^2}$ for $(x, y) \ne (0, 0)$, and $f(0, 0) = 0$, then for each x, the single variable function h given by $h(y) = f(x, y)$ is continuous. Also, for each y, the single variable function g given by $g(x) := f(x, y)$ is continuous. Nevertheless, f is not continuous in $(0, 0)$.

See the homework for an even more striking example of a similar nature.

Open sets; Boundary:

Sometimes it comes convenient to establish some language to describe certain geometric features of sets in the plane or space (or $\mathbb{R}^n$).

We call a set *open* if, with each point P, it contains also a small ball around it. (We often refer to this ball as a 'neighborhood of P'.) More formally $\mathbb{M} \subseteq \mathbb{R}^n$ is called an open set, if for every $P \in \mathbb{M}$, there is a (small) $\delta > 0$ such that the entire 'ball' $B_\delta(P) := \{Q \mid \|\vec{PQ}\| < \delta\}$ is contained in $\mathbb{M}$. For instance the set $\{(x, y) \mid y > 0\}$ is open, but the set $\{(x, y) \mid y \geq 0\}$ is not open.

We say a point P in on the boundary of a set $\mathbb{M}$, if every ball $B_\delta(P)$ contains both points from $\mathbb{M}$ and points that are not in $\mathbb{M}$. We call a set closed, if it contains all its boundary points.

Despite an impression created by these words, sets can possibly be neither open nor closed. $\mathbb{R}^n$ itself is an example of a set that is both open and closed.