Continuity

Alex Nita

Abstract

In this section we try to get a very rough handle on what's happening to a function f in the neighborhood of a point P. If I have a function $f : \mathbb{R}^n \to \mathbb{R}^m$ one of the first things I would want to check is it's continuity at P, because then at least I'd know if my function, which may represent some physical quantity, is "well behaved" or "moody"/"shifty" there. It's a first approximation, a sort of "eyballing" to see what's happening, roughly. To figure this out, I need to know what's happening to f *near* P. If I were to summarize the main idea behind continuity, I'd say the continuity of f at P is entirely to do with f being *stable* near P, not jumping or oscillating wildy, not varying erratically, basically being somehow *contained* or *controlled* near P. Continuity is about f being stabilized near P, about it's outputs, the $\mathbf{y}$-values, staying within any given neighborhood whenever its inputs, the $\mathbf{x}$-values, stay within a prescirbed neighborhood of P. The emphasis of continuity at P is on what's happening *around* P, not *at* P. There is a nice visual in terms of the graph of f: continuity at P means the graph of f near P doesn't tear or crinckle horribly–it has a kind of consistency near P. In this section we try to make all this a bit more rigorous and computationally useful.

How you should read this: Firstly, I have included several interesting examples of vector-valued functions. You should take a look at those examples if you're wondering why we bother with all this vector-valued business. It's because there are inherently interesting functions with lots of physical and geometric content. But don't get hung up on the examples, just read them at leisure to see some interesting motivation. Secondly, I have included the technical definitions of limit and continuity. Don't get hung up on those, or the examples which illustrate their use. They're there as a supplement to the book, and for those curious to see the underlying rigor to all this. The really important thing here, in my opinion, is the theorem which says that any vector-valued function $f = (f_1, \dots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ is continous at a point $\mathbf{a} \in \mathbb{R}^n$ if and only if each of it's m component functions $f_i : \mathbb{R}^n \to \mathbb{R}$ is continuous at $\mathbf{a}$. For this shows that the continuity behavior of vector-valued functions is totally determined by the behavior of its real-valued components. Hence, we only ever need to worry about continuity of real-valued functions $f : \mathbb{R}^n \to \mathbb{R}$. This is why most calculus books only ever deal with this case. But don't worry about the proof–I haven't even included it, because it's a simple chasing of εs and δs. It's technical and not worth your while in this course. Finally, look at the examples at the end. The reason for all the fuss over limits and continuity is because if you look at the important examples of functions we are likely to encounter–coordinate transformations, physical examples, etc.–we'd really like to know whether they behave erratically anywhere, and if so why. It's an unfortunate fact of life that not all functions are well-behaved in this sense. But what about our canonical coordinate transformations and the physical examples we are likely to encounter? It would be nice to know that at least *they* behave well! Let's see if they do.

One more theorem is worth noting: the one that says sums, products, scalar multiples and compositions of continuous functions are continuous. For knowing this gives us a way to check whether a given function is continuous by merely checking the simple functions of which it is made up. Thus, knowing that sine and polynomials are continous, we know that $2\sin(xy^2) + 5x^5y^4$ is continuous.

I have provided references to books you can consult for proofs and further discussion on these topics. Some of these books are standard undergraduate analysis or linear algebra texts, others you may never see in such a course, at least I haven't. In any case, they're there for those of you interested in seeing more of what's under the hood.

1 Examples of Vector-Valued Functions

Consider a function $f : \mathbb{R}^n \to \mathbb{R}^m$. We have already encountered a large collection such functions in the notes on vectors and matrices, the linear functions, which are of the form $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ (here A is an $m \times n$ matrix and $\mathbf{x}$ and $\mathbf{b}$ are considered as "vertical" or "column vectors", in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively). Now we consider the general case, with f not necessarily linear. We *think* of such a function as **vector-valued**, although it is just $\mathbf{y}$-valued, where $\mathbf{y}$ isn't a real number but an m-tuple of real numbers $\mathbf{y} = (y_1, \ldots, y_m)$, a point in $\mathbb{R}^m$. What f does is send points $\mathbf{x} = (x_1, \ldots, x_n)$ to points $\mathbf{y} = (y_1, \ldots, y_m)$, but how is each of the real entries $y_1, \ldots, y_m$ in $\mathbf{y}$ determined? Well, each y_i is a function of $\mathbf{x}$, in this case a *real-valued* function of $\mathbf{x}$, which we call the i**th component function of** f, and denote it f_i,

$$\begin{aligned} f_i &: \mathbb{R}^n \to \mathbb{R} \\ y_i &= f_i(\mathbf{x}) \end{aligned} \tag{1.1}$$

This means

$$\big(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\big) = (y_1, \ldots, y_n) = \mathbf{y} = f(\mathbf{x}) \tag{1.2}$$

This can be concisely written

$$\boxed{f = (f_1, \ldots, f_m)} \tag{1.3}$$

Remark 1.1 *The following examples are explained in much greater geometric detail in a separate section on coordinates, and you should refer to that for their geometric content and derivations. For now, we just consider the functions in the abstract.* ■

Example 1.2 *A simple example of such a function is a change-of-coordinates. Consider the change-of-coordinates map $f : \mathbb{R}^3 \to \mathbb{R}^3$ changing spherical to Cartesian coordinates:*

$$\boxed{\begin{aligned} &f : \mathbb{R}^3 \to \mathbb{R}^3 \\ &(x, y, z) = f(\rho, \theta, \varphi) = \big(\rho \sin\varphi \cos\theta,\ \rho \sin\varphi \sin\theta,\ \rho\cos\varphi\big) \end{aligned}}$$

The component functions$f_1, f_2, f_3 : \mathbb{R}^3 \to \mathbb{R}$ of f are

$$\begin{aligned} f_1(\rho, \theta, \varphi) &= \rho \sin\varphi \cos\theta = x \\ f_2(\rho, \theta, \varphi) &= \rho \sin\varphi \sin\theta = y \\ f_3(\rho, \theta, \varphi) &= \rho \cos\varphi = z \end{aligned}$$

■

Example 1.3 *Consider the rotation of the plane $\mathbb{R}^2$ through a fixed angle θ (rotated coordinates):*

$$\boxed{\begin{aligned} &f : \mathbb{R}^2 \to \mathbb{R}^2 \\ &(x', y') = f(x, y) = \big(x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta\big) \end{aligned}}$$

For example, if we fix $\theta = \frac{\pi}{6}$, then the new coordinates are

$$(x', y') = f(x, y) = \left(\frac{\sqrt{3}x}{2} - \frac{y}{2}, \frac{x}{2} + \frac{\sqrt{3}y}{2}\right)$$

The component functions $f_1, f_2 : \mathbb{R}^2 \to \mathbb{R}$ are

$$f_1(x,y) = \frac{\sqrt{3}x}{2} - \frac{y}{2} = x'$$
$$f_2(x,y) = \frac{x}{2} + \frac{\sqrt{3}y}{2} = y'$$

This is actually an example of a linear *function, for we can write it in matrix form:*

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

■

Example 1.4 *Suppose we rotate not the xy-plane, but some arbitrary plane $ax+by+cz+d = 0$ in $\mathbb{R}^3$ through a fixed angle θ. The function $R : \mathbb{R}^3 \to \mathbb{R}^3$ for this is dubbed **Rodriguez' rotation formula**: take the normal vector $\mathbf{n} = \langle a, b, c\rangle$ to the plane, and construct the following function $f : \mathbb{R}^3 \to \mathbb{R}^3$:*

$$f(\mathbf{x}) = \mathbf{n} \times \mathbf{x}$$

That is,

$$f(x,y,z) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a & b & c \\ x & y & z \end{vmatrix} = (bz - cy,\ xc - az,\ ay - bx)$$

This is a linear function, and has matrix representation.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

The matrix above,

$$A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}$$

is called the "cross-product matrix", in light of the fact that it was gotten from the cross-product function f. Using this matrix A, Rodriguez' rotation formula then gives the rotation function $R : \mathbb{R}^3 \to \mathbb{R}^3$ in terms of the cross-product matrix above,

$$R : \mathbb{R}^3 \to \mathbb{R}^3$$
$$R(\mathbf{x}) = I\mathbf{x} + \sin\theta A\mathbf{x} + (1 - \cos\theta)A^2\mathbf{x}$$

(We think of $\mathbf{x}$ as a column vector, and I is the 3×3 identity matrix with 1s along the diagonal and 0s everywhere else.) It is a fact of linear algebra that there are coordinates $(u, v) = u\mathbf{r}+v\mathbf{s}$ for the plane $ax + by + cz + d = 0$ in which the function R has the obvious rotation form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

Of course, we take the third coordinate vector for $\mathbb{R}^3$ to be $\mathbf{n}$ itself, so that any point in this coordinate system is given by $(t, u, v) = t\mathbf{n} + u\mathbf{r} + v\mathbf{s}$. ■

Example 1.5 *Recall the rotation-followed-by-a-translation function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given in matrix form by*

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$$

This function was described in greater detail in the lecture notes on vectors and matrices. ■

Example 1.6 *Recall the function $\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}^2$ which reflects the entire plane across a given line $\ell = \{\lambda\mathbf{v} \mid \lambda \in \mathbb{R}\}$, where $\mathbf{v} = \langle a, b\rangle$ is some given unit vector sitting on the line. It is given by*

$$\mathcal{R}(\mathbf{x}) = 2(\mathbf{x}\cdot\mathbf{v})\mathbf{v} - \mathbf{x}$$

It is in fact a linear function and can be described by a matrix:

$$\mathcal{R}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2a^2 - 1 & 2ab \\ 2ab & 2b^2 - 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

where we write points/vectors as column vectors. This function, too, was described in greater detail in the lecture notes on vectors and matrices. ■

Example 1.7 (Electric and Magnetic Fields) *Consider the functions*

$$\mathbf{E} : \mathbb{R}^4 \to \mathbb{R}^3$$
$$\mathbf{B} : \mathbb{R}^4 \to \mathbb{R}^3$$

describing, respectively, the ***electric*** *and* ***magnetic vector fields*** *in $\mathbb{R}^3$, varying over time,*

$$\mathbf{E}(\mathbf{x}, t) = \mathbf{E}(x, y, z, t)$$
$$\mathbf{B}(\mathbf{x}, t) = \mathbf{B}(x, y, z, t)$$

giving, at each point $\mathbf{x} = (x, y, z)$, the magnitude and direction of the electric, resp. magnetic, fields due to a charged source. If the fields $\mathbf{E}$ and $\mathbf{B}$ are constant, i.e. unchanging over time, then the time variable can be dropped and we can consider them as functions from $\mathbb{R}^3$ to $\mathbb{R}^3$, in which case we call them, respectively, the ***electrostatic*** *and* ***magnetostatic fields****.*

The electric field due to a single point-charge q_1 has a rather simple form, due to Coulomb's Law:

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\frac{q_1(\mathbf{r} - \mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|^3} = \frac{1}{4\pi\varepsilon_0}\frac{q_1}{|\mathbf{r} - \mathbf{r}_1|^2}\widehat{\mathbf{r} - \mathbf{r}_1}$$

where $\mathbf{r}_1$ is the position of the source charge q_1 and $\mathbf{r}$ is the point of interest in $\mathbb{R}^3$ a distance $|\mathbf{r} - \mathbf{r}_1|$ away from q_1 (note, therefore, the radial symmetry of $\mathbf{E}$), with $\widehat{\mathbf{r} - \mathbf{r}_1} = \frac{\mathbf{r}-\mathbf{r}_1}{|\mathbf{r}-\mathbf{r}_1|}$ the unit vector in the direction of q. ε_0 is a constant, called the ***vacuum permitivity****. We'd like to determine the electric field created by q_1 at $\mathbf{r} \in \mathbb{R}^3$, for then maybe we would know what would happen to another charged particle q at that point. By Coulomb's law, if we were to drop that second particle at $\mathbf{r}$, the the* ***electrostatic force*** *exerted on q at position $\mathbf{r}$ (assuming $\mathbb{R}^3$ is a vacuum) would be*

$$\mathbf{F}(\mathbf{r}) = q\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\frac{qq_1(\mathbf{r} - \mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|^3}$$

This defines another vector-valued function,

$$\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$$

If the electric field evolves over time, and so is not static, then of course so does $\mathbf{F}$*, in which case we should define* $\mathbf{F} : \mathbb{R}^4 \to \mathbb{R}^3$*.*

Since the electric fields satisfy the ***superposition principle****, which says that the electric fields* $\mathbf{E}_1$ *and* $\mathbf{E}_2$ *due to source charge particles* q_1 *and* q_2 *are additive, we get for any fixed number of charged particles* $q_1, \ldots, q_n$ *sitting at positions* $\mathbf{r}_1, \ldots, \mathbf{r}_n$*, respectively, that the cumulative electric field created by* $q_1, \ldots, q_n$ *is*

$$\mathbf{E}(\mathbf{r}) = \sum_{i=1}^{n} \mathbf{E}_i = \frac{1}{4\pi\varepsilon_0} \sum_{i=1}^{n} \frac{q_i(\mathbf{r} - \mathbf{r}_i)}{|\mathbf{r} - \mathbf{r}_i|^3}$$

We next pretend that a charged surface S *is a continuous medium consisting of uncountably many such source charges* q_i*, with possibly different charges, depending of course on their location* $\mathbf{x}$ *in* S*, but now we change our thinking slightly and view the point-particle* q_i *sitting at* $\mathbf{x}$ *as having not a charge, but a charge* ***charge density****, denoted* $\rho(\mathbf{x})$*. This is basically an infinitessimal version of charge, namely a charge per unit volume. From this, and with the assumption that* ρ *varies continuously (so there aren't two points in* S *sitting close to each other with wildly varying charge densities), we get a continuous analog of the above summation forumla via limits of Riemann sums,*

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \int_S \frac{\rho(\mathbf{x})(\mathbf{r} - \mathbf{x})}{|\mathbf{r} - \mathbf{x}|^3} d\mathbf{x}$$

for the electric field at any point $\mathbf{r}$ *outside of* S*.*

Of course, determining $\mathbf{E}$*,* $\mathbf{F}$ *and* $\mathbf{B}$ *for particular configurations of charges in a surface* S *is the hard part. It is the main task of the physicist. All we know about these functions is the constraints they satisfy (in a vacuum), which were first spelled out by James Clerk Maxwell in 1861-1862, and are now known as* ***Maxwell's equations****:*

$$\begin{aligned}
\nabla \cdot \mathbf{E} &= 0 \\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \times \mathbf{B} &= \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}.
\end{aligned}$$

■

Example 1.8 (Newtonian Gravitation) *A similar situation is to be found with Newtonian, or classical, gravitation. The* ***gravitational field*** *created by a point-particle with mass* m_1 *is a vector-valued function*

$$\begin{aligned}
&\mathbf{g} : \mathbb{R}^3 \to \mathbb{R}^3 \\
&\mathbf{g}(\mathbf{r}) = -\frac{Gm_1(\mathbf{r} - \mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|^3} = -\frac{Gm_1}{|\mathbf{r} - \mathbf{r}_1|^2} \widehat{\mathbf{r} - \mathbf{r}_1}
\end{aligned}$$

where $\mathbf{r}$ *is the point in* $\mathbb{R}^3$ *a distance* $|\mathbf{r} - \mathbf{r}_1|$ *from our source particle (hence the radial symmetry of* $\mathbf{g}$*), and* G *is the* ***gravitational constant****. If a second particle of mass* m *were to be dropped at the point, then* ***Newton's law of universal gravitation*** *says that the force exerted on the second particle at the point* $\mathbf{r}$ *by the first is*

$$\begin{aligned}
&\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3 \\
&\mathbf{F}(\mathbf{r}) = m\mathbf{g}(\mathbf{r}) = -\frac{Gmm_1}{|\mathbf{r} - \mathbf{r}_1|^2} \widehat{\mathbf{r} - \mathbf{r}_1}
\end{aligned}$$

If the second particle of mass m moves along a trajectory over time, then $\mathbf{r}$ *gives the position of that particle at time t, and so defines a path in $\mathbb{R}^3$:*

$$\mathbf{r} : \mathbb{R} \to \mathbb{R}^3, \qquad \mathbf{r} = \mathbf{r}(t)$$

In this scenario, Newton's second law *of motion says that the force* $\mathbf{F}$ *exerted on the second particle by the first must satisfy a first order differential equation:*

$$\mathbf{F} = m\mathbf{a} = \frac{d^2\mathbf{r}(t)}{dt^2}$$

assuming, of course, that the path of the particle is smooth enough to be differentiated. If there are n particles of respective masses $m_1, \ldots, m_n$, then the above considerations apply pairwise to any two of them, m_i and m_j. But what about the whole system of n particles? If the particles are held fixed, then as with the electrostatic field the cumulative graviational field on an outside particle of mass m is the sum of these,

$$\mathbf{F} = \sum_{i=1}^{n} \mathbf{F}_i = \sum_{i=1}^{n} -\frac{Gmm_i}{|\mathbf{r} - \mathbf{r}_i|^2}\widehat{\mathbf{r} - \mathbf{r}_i}$$

But if the particles are not *held fixed, but interact with each other (like planets in a solar system, for example), then we get from the law of universal gravitation and the second law of motion that they must simultaneously satisfy the n equations*

$$m_i \frac{d^2\mathbf{r}_i(t)}{dt^2} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{Gm_im_j}{|\mathbf{r}_j(t) - \mathbf{r}_i(t)|^2} \mathbf{r}_j\widehat{(t) - \mathbf{r}_i}(t)$$

*If the n particles were all the planets, moons, and sun in our solar system, for example, we would consider ourselves successful if we could predict their paths accurately–accurately enough, say, to predict solar or lunar eclipses. That would mean having explicit formulas for the paths $\mathbf{r}_i(t)$ of each planet, moon or the sun, for then we could simply solve for t in any particular configuration of the system, say with the moon exactly in between earth and the sun, or with the earth exactly in between the sun and the moon. Finding the paths $\mathbf{r}_i(t)$, then, is the task, and this task is notoriously difficult, and has become known as the **n-body problem**. The case $n = 2$ has been solved, and partial results are in for the case $n = 3$, but for $n > 3$ the problem is still open. For a quick overview, see `http://en.wikipedia.org/wiki/N-body_problem`.* ■

2 Continuity of a Vector-Valued Function at a Point

2.1 Preliminaries: The Triangle and Revese Triangle Inequalities

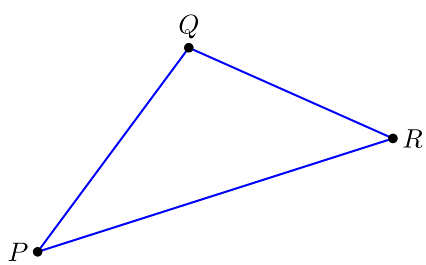
Let us recall the definition of *distance* in $\mathbb{R}^n$: the distance between two points $P = (x_1, \ldots, x_n)$ and $Q = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ is given by an analog of the Pythagorean theorem:

$$d(P, Q) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + (y_n - x_n)^2} \tag{2.1}$$

In $\mathbb{R}^2$ this is *exactly* the Pythagorean theorem. Alternatively, we can view this as the *length/magnitude* of the displacement vector $\overrightarrow{PQ}$ from P to Q,

$$\begin{aligned} |\overrightarrow{PQ}| &= |\vec{Q} - \vec{P}| \\ &= |\langle y_1, \ldots, y_n \rangle - \langle x_1, \ldots, x_n \rangle| \\ &= |\langle y_1 - x_1, \ldots, y_n - x_n \rangle| \\ &= \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + (y_n - x_n)^2} \end{aligned} \tag{2.2}$$

There is an important inequality, the **triangle inequality**, relating the distances between points P, Q and R: namely, the shortest distance between P and R is the straight-line distance, shorter than the distance from P to Q plus the distance from Q to R,



$$d(P, R) \le d(P, Q) + d(Q, R)$$

or

$$|\overrightarrow{PR}| \le |\overrightarrow{PQ}| + |\overrightarrow{QR}|$$

Why is this important? It's our main tool for *controlling* f near P. Inequalities are great for control. Before I explain all this, let me first prove the triangle inequality. The first step toward it's proof requires another, simpler inequality, the **Cauchy-Schwartz inequality**. We remark only that it's easier to work with vectors $\vec{x}$, $\vec{y}$ and $\vec{z}$ than points and their displacement vectors, because we can always go back and forth between points and vectors anyway. It makes the proofs easier to work with $\vec{x}$ rather than $\overline{PQ}$, etc.

Theorem 2.1 (Cauchy-Schwartz Inequality) *For any vectors* $\vec{x}, \vec{y} \in \mathbb{R}^n$ *we have*

$$\boxed{|\vec{x} \cdot \vec{y}| \le |\vec{x}||\vec{y}|} \tag{2.3}$$

Proof: Note that in $\mathbb{R}^2$ and $\mathbb{R}^3$ this follows from the fact that we have well-defined angles between vectors: this supplies the identity $\vec{x} \cdot \vec{y} = |\vec{x}||\vec{y}| \cos\theta$, and along with the fact that $|\cos\theta| \le 1$ we get our inequality. But in general, unless we define angles somehow,[1] we cannot use this fact, but anyway we don't need to use it–there's a direct way to prove this without angles. Define the *nonnegative* real-valued function of t

$$f(t) = |\vec{x} - t\vec{y}|^2$$

[1] This *can* actually be done, using the dot product, in fact: simply let $\theta = \cos^{-1}\left(\frac{\vec{x} \cdot \vec{y}}{|\vec{x}||\vec{y}|}\right)$. Cf. http://en.wikipedia.org/wiki/Euclidean_space#Angle.

Then, foiling the right-hand out using the distributive and scalar properties of the dot product we get

$$f(t) = |\vec{x} - t\vec{y}|^2 = (\vec{x} - t\vec{y}) \cdot (\vec{x} - t\vec{y}) = \vec{x} \cdot \vec{x} - 2t\vec{x} \cdot \vec{y} + t^2 \vec{y} \cdot \vec{y}$$

which shows that f is a parabola sitting above the t-axis:

$$f(t) = at^2 + bt + c \geq 0$$

where $a = \vec{y} \cdot \vec{y} = |\vec{y}|^2$, $b = -2\vec{x} \cdot \vec{y}$ and $c = \vec{x} \cdot \vec{x} = |\vec{x}|^2$. Since f is sitting on or above the t-axis, it has at most one real double root, or else two imaginary roots. This means, when we use the quadratic formula $t = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ for the roots of f, the discriminant $b^2 - 4ac$ is either 0 or negative, i.e.

$$b^2 - 4ac \leq 0$$

Plugging our a, b and c into this inequality gives

$$(-2\vec{x} \cdot \vec{y})^2 - 4|\vec{y}|^2|\vec{x}|^2 \leq 0$$

which means

$$\not{4}|\vec{x} \cdot \vec{y}|^2 \leq \not{4}|\vec{y}|^2|\vec{x}|^2$$

Taking the square root of both sides gives the result. ■

Theorem 2.2 (Triangle Inequality) *For any two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have*

$$\boxed{|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|} \tag{2.4}$$

Proof: By the Cauchy-Schwartz inequality we have

$$\begin{aligned} |\vec{x} + \vec{y}|^2 &= (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) \\ &= \vec{x} \cdot \vec{x} + 2\vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y} \\ &\leq \vec{x} \cdot \vec{x} + 2|\vec{x} \cdot \vec{y}| + \vec{y} \cdot \vec{y} \qquad \text{because } a \leq |a| \text{ always} \\ &\leq |\vec{x}|^2 + 2|\vec{x}||\vec{y}| + |\vec{y}|^2 \qquad \text{by the Cauchy-Schwartz inequality} \\ &= (|\vec{x}| + |\vec{y}|)^2 \end{aligned}$$

and taking the square root gives the triangle inequality. ■

Let us now prove the **reverse triangle inequality** using the triangle inequality, which is really the inequality we need for continuity.

Corollary 2.3 (Reverse Triangle Inequality) *For any thwo vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have*

$$\boxed{\big||\vec{x}| - |\vec{y}|\big| \leq |\vec{x} - \vec{y}|} \tag{2.5}$$

Proof: By the triangle inequality we get the following two inequalities:

$$\begin{aligned} |\vec{x}| = |\vec{x} - \vec{y} + \vec{y}| \leq |\vec{x} - \vec{y}| + |\vec{y}| \\ |\vec{y}| = |\vec{y} - \vec{x} + \vec{x}| \leq |\vec{y} - \vec{x}| + |\vec{x}| \end{aligned}$$

Since $|\vec{y} - \vec{x}| = |\vec{x} - \vec{y}|$ (because in general we have $|a\vec{v}| = |a||\vec{v}|$, so with $a = -1$ we get $|-\vec{v}| = |\vec{v}|$) we have that

$$\begin{aligned} |\vec{x}| \leq |\vec{x} - \vec{y}| + |\vec{y}| \\ |\vec{y}| \leq |\vec{x} - \vec{y}| + |\vec{x}| \end{aligned}$$

Subtracting $|\vec{y}|$ from both sides of the first, and $|\vec{x}|$ from both sides of the second, we get that

$$|\vec{x}| - |\vec{y}| \leq |\vec{x} - \vec{y}|$$
$$-(|\vec{x}| - |\vec{y}|) \leq |\vec{x} - \vec{y}|$$

If we multiply the second inequality above by -1, we get that

$$|\vec{x}| - |\vec{y}| \geq -|\vec{x} - \vec{y}|$$

Combining the first of the above two inequalities with this one gives

$$-|\vec{x} - \vec{y}| \leq |\vec{x}| - |\vec{y}| \leq |\vec{x} - \vec{y}|$$

which is precisely the statement of the reverse triangle inequality. (The statement $|a| \leq b$ is equivalent to $-b \leq a \leq b$ for all real numbers a and b. For certainly whenever $|a| \leq b$ we have $a \leq |a| \leq b$, and so $-b \leq -|a| \leq a$, too, i.e. $-b \leq a \leq b$. And conversely, if $-b \leq a \leq b$, then since either $a = |a|$ or $a = -|a|$ depending on whether a is positive or not: in the first case, we have $-b \leq a = |a| \leq b$, so in particular $|a| \leq b$, while in the second case we have $-b \leq a = -|a| \leq b$, so in particular $-b \leq -|a|$, or $|a| \leq b$.) ■

Example 2.4 *Consider the vectors* $\vec{x} = \langle 1, 2, 3 \rangle$ *and* $\vec{y} = \langle -2, 0, 4 \rangle$ *in* $\mathbb{R}^3$. *Let us verify all three inequalities. First, the Cauchy-Schwartz inequality: since* $100 \leq 280$ *implies* $10 \leq \sqrt{280}$, *we have*

$$\begin{aligned} |\vec{x} \cdot \vec{y}| &= |\langle 1, 2, 3 \rangle \cdot \langle -2, 0, 4 \rangle| \\ &= |-2 + 0 + 12| \\ &= 10 \\ &\leq \sqrt{280} \\ &= \sqrt{14}\sqrt{20} \\ &= |\langle 1, 2, 3, \rangle||\langle -2, 0, 4 \rangle| \\ &= |\vec{x}||\vec{y}| \end{aligned}$$

Next, the triangle inequality: because

$$\begin{aligned} 54 = 14 + 20 + 20 &= (\sqrt{14})^2 + 2 \cdot 10 + (\sqrt{20})^2 \\ &\leq (\sqrt{14})^2 + 2 \cdot \sqrt{280} + (\sqrt{20})^2 = (\sqrt{14} + \sqrt{20})^2 \end{aligned}$$

we have $\sqrt{54} \leq \sqrt{14} + \sqrt{20}$, *and so*

$$\begin{aligned} |\vec{x} + \vec{y}| &= |\langle 1, 2, 3 \rangle + \langle -2, 0, 4 \rangle| \\ &= |\langle -1, 2, 7 \rangle| \\ &= \sqrt{54} \\ &\leq \sqrt{14} + \sqrt{20} \\ &= |\langle 1, 2, 3 \rangle| + |\langle -2, 0, 4 \rangle| \\ &= |\vec{x}| + |\vec{y}| \end{aligned}$$

Finally, the reverse triangle inequality: Since $20 \leq 56 = 4 \cdot 14$, *we have* $\sqrt{20} \leq 2\sqrt{14}$ *and* *therefore* $\sqrt{20} - \sqrt{14} \leq \sqrt{14}$, *which means*

$$\begin{aligned} \big||\vec{x}| - |\vec{y}|\big| &= \big||\langle 1, 2, 3\rangle| - |\langle -2, 0, 4\rangle|\big| \\ &= \big|\sqrt{14} - \sqrt{20}\big| \\ &= \sqrt{20} - \sqrt{14} \\ &\leq \sqrt{14} \\ &= |\langle 3, 2, -1\rangle| \\ &= |\langle 1, 2, 3\rangle - \langle -2, 0, 4\rangle| \\ &= |\vec{x} - \vec{y}| \end{aligned}$$

■

2.2 Limits and Continuity of Real- and Vector-Valued Functions

Let's just drop the limit/continuity bomb and deal with the carnage afterward, i.e. let's just give the dry definition and try to explain its meaning to humans afterward. We say that a vector-valued function $f : \mathbb{R}^n \to \mathbb{R}^m$ has a **limit** in $\mathbb{R}^m$, which we denote $\mathbf{L} = (L_1, \ldots, L_m)$, at a point $\mathbf{a} = (a_1, \ldots, a_n)$ in $\mathbb{R}^n$, and we write

$$\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = \mathbf{L} \tag{2.6}$$

if for all $\varepsilon > 0$ we can find a $\delta > 0$ such that for all $\mathbf{x}$ different from $\mathbf{a}$ (but near $\mathbf{a}$), the following property is satisfied:

$$\text{If } d(\mathbf{x}, \mathbf{a}) < \delta, \text{ then } d\big(f(\mathbf{x}), \mathbf{L})\big) < \varepsilon. \tag{2.7}$$

We can rephrase property (2.7) using equation (2.2) as

$$\text{If } |\mathbf{a} - \mathbf{x}| < \delta, \text{ then } |\mathbf{L} - f(\mathbf{x})| < \varepsilon. \tag{2.8}$$

Remark 2.5 *Notice that here* $\mathbf{a}$ *and* $\mathbf{x}$ *are points, but we treated them like vectors and subtracted them and took their magnitude. We could have been pedantic and written* $\vec{\mathbf{a}}$ *and* $\vec{\mathbf{x}}$ *every time we did this, but let's face it, that would just clutter notation and get us all anxious about whether we've forgotten an arrow or forgot to add an arrow. The larger issue in the distinction between points and vectors is actually not this, not the pedantry, but rather a fundamental distinction of what space we're working with. If we consider the points as lying on a fixed surface* S *or curve* C *or something analogous, then they should definitely stay points, because adding them might remove their sum from the surface or curve (for example adding the point* $(0,1)$ *to the point* $(0,-1)$ *gives the point* $(0,0)$*. If we were working on the circle* $x^2 + y^2 = 1$*, then their sum would not be on the circle–problem!). But if we're just talking about* $\mathbb{R}^n$ *this isn't a problem, for adding them keeps them in* $\mathbb{R}^n$*. If we get too hung up on the distinction between points and vectors we run the risk of trying to do Riemannian geometry instead of Calc 3! Thus, for the remainder of these notes, and for the class, let us try not to worry too much about the distinction between points and vectors and just treat them as interchangeable unless the context demands our attention to their difference.* ■

Next, we say that $f : \mathbb{R}^n \to \mathbb{R}^m$ is **continuous at a point** $\mathbf{a} = (a_1, \ldots, a_n)$ in $\mathbb{R}^n$ if f has a limit $\mathbf{L}$ there and moreover that limit equals the value of f at $\mathbf{a}$, $\mathbf{L} = f(\mathbf{a})$. We can rephrase all of this in terms of ε's and δ's, if we wanted to, of course. If f is continuous at $\mathbf{a}$, then we write

$$\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = f(\mathbf{a}) \tag{2.9}$$

Remark 2.6 *OK, now let us assess the damage, let us analyze these definitions. First of all, notice that we have complete freedom in choosing our positive number* ε*. The idea here is that in the range of* f *we can make the distance between* $f(\mathbf{x})$ *to* $\mathbf{L}$ *as small as we want, provided back in the domain of* f *we're within* δ *of our point* $\mathbf{a}$*, for* some δ*. That is, you give me a distance you want* f *to stay within from* $\mathbf{L}$*, and I can find you a neighborhood of* $\mathbf{a}$ *which is* <u>*entirely*</u> *mapped to within* ε *of* $\mathbf{L}$ *by* f*. The key point here is that the neighborhood of* $\mathbf{a}$ *may change depending on how small you want to keep the values of* f *from* $\mathbf{L}$*. If you make* ε *smaller, you'll probably have to shrink* δ *and so the neighborhood of* $\mathbf{a}$*.*

The picture you should have in mind is of a large lever. The lever is the function, the handle is the domain of the function, and the end on the other side of the fulcrum is the range of the *function. If you push the lever a certain amount, you raise an object another given amount. The idea of continuity in such a situation is intuitive, namely the lever can't disappear at one place and reappear at another, and the way we formalize this intuition is by saying that if you want to look really really close to a point near the object you're trying to lift (only ε far), you'll see the end of the lever there so long as you only wiggle the handle no more than a given amount (your δ-neighborhood). There can be no skipping or disappearing and reappearing in such a situation, because if something like that happened, we could shrink our ε-window to a distance smaller than the size of the skip, and then the end of the lever would disappear if we wiggle the handle at all.*

There is another benefit to the formalism, namely that it allows you to show when a given function actually is continuous. Typically that sort of exercise is to be carried out in an Undergraduate Analysis course, not here. But we gave the definition here for the record, and to prove one or two results in case anybody's interested in seeing how this stuff works.

Yet another benefit of this formalism is that it illustrates the philosophy of calculus quite clearly: What's important is not what happens at a point $\mathbf{a}$*, but near a point* $\mathbf{a}$. ■

Remark 2.7 *Let us look at the case where $m = 1$, i.e. the case of a real-valued function $f : \mathbb{R}^n \to \mathbb{R}$. In the case of $\mathbb{R}$, the magnitude is simply the absolute value, for if $\mathbf{x} = (x_1)$, then $|\mathbf{x}| = |(x_1)| = \sqrt{x_1^2} = |x_1|$.* ■

Example 2.8 *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^2 + y^2$. I claim it is continuous at any point (a, b) and can show this directly by applying the reverse triangle inequality:*

$$|f(x,y) - f(a,b)| = \big|(x^2+y^2) - (a^2+b^2)\big| = \big||(x,y)| - |(a,b)|\big| \le \big|(x,y) - (a,b)\big|$$

so that for any $\varepsilon > 0$ we can just choose $\delta = \varepsilon$, and then we get that $\big|(x,y) - (a,b)\big| < \delta = \varepsilon$ implies $|f(x,y) - f(a,b)| < \varepsilon$. ■

Example 2.9 *Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be given by $f(x, y, z) = (2x - 3y + z,\ -x + y + 2z)$. Let us show that f is continuous at every point $\mathbf{a} = (a, b, c)$ in $\mathbb{R}^3$. Toward this end, let us note that for any $\mathbf{v} = (u, v, w)$ we have*

$$\begin{aligned}
|f(\mathbf{v})| &= |f(u,v,w)| \\
&= |(2u - 3v + w,\ -u + v + 2w)| \\
&= \sqrt{(2u-3v+w)^2 + (-u+v+2w)^2} \\
&= \sqrt{\Big[(2,-3,1)\cdot(u,v,w)\Big]^2 + \Big[(-1,1,2)\cdot(u,v,w)\Big]^2} \\
&\le \sqrt{\Big[|(2,-3,1)||(u,v,w)|\Big]^2 + \Big[|(-1,1,2)||(u,v,w)|\Big]^2} \qquad \textit{Cauchy-Schwartz ineq.} \\
&= \sqrt{|(2,-3,1)|^2|(u,v,w)|^2 + |(-1,1,2)|^2|(u,v,w)|^2} \\
&= \sqrt{\big[|(2,-3,1)|^2 + |(-1,1,2)|^2\big]|(u,v,w)|^2} \\
&= \sqrt{|(2,-3,1)|^2 + |(-1,1,2)|^2}\sqrt{|(u,v,w)|^2} \\
&= \sqrt{(4+9+1) + (1+1+4)}|(u,v,w)| \\
&= \sqrt{20}|\mathbf{v}|
\end{aligned}$$

Consequently, since $f(\mathbf{x}) - f(\mathbf{a}) = f(\mathbf{x} - \mathbf{a})$, i.e. $f(x, y, z) - f(a, b, c) = f(x - a, y - b, z - c)$ (check this!), we have

$$|f(\mathbf{x}) - f(\mathbf{a})| = |f(\mathbf{x} - \mathbf{a})| \le \sqrt{20}|\mathbf{x} - \mathbf{a}|$$

Thus, for any $\varepsilon > 0$ that we pick we can find our $\delta > 0$, namely $\delta = \frac{\varepsilon}{\sqrt{20}}$, for then, whenever $|\mathbf{x} - \mathbf{a}| < \delta$, we have

$$|f(\mathbf{x}) - f(\mathbf{a})| \leq \sqrt{20}|\mathbf{x} - \mathbf{a}| < \sqrt{20}\delta = \sqrt{20}\frac{\varepsilon}{\sqrt{20}} = \varepsilon$$

Voila! We have shown that $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$ for any $\mathbf{a} = (a, b, c)$ in $\mathbb{R}^3$. ∎

Remark 2.10 *This example is a lot harder than what you are asked to do on homeworks or tests, but the method used in it is universally applicable. The basic idea is to try get an inequality of the form*

$$|f(\mathbf{x}) - f(\mathbf{a})| \leq C|\mathbf{x} - \mathbf{a}|$$

for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ for some positive constant C. If you get to this point, you're home free (but beware, it's not always possible; a good example is the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 1/x$). The intuition is clear: let $\mathbf{x}$ approach $\mathbf{a}$, then you squeeze $|\mathbf{x} - \mathbf{a}|$, which consequently squeezes $|f(\mathbf{x}) - f(\mathbf{a})|$ by a proportionate amount. Formally, say you want to squeeze $|f(\mathbf{x}) - f(\mathbf{a})|$ to under ε, then all you have to do is squeeze $|\mathbf{x} - \mathbf{a}|$ to under ε/C.

As a matter of fact, if you get an inequality of the type $|f(\mathbf{x}) - f(\mathbf{a})| \leq C|\mathbf{x} - \mathbf{a}|$, you basically succeeded in showing continuity. You can just say

$$\boxed{|f(\mathbf{x}) - f(\mathbf{a})| \leq C|\mathbf{x} - \mathbf{a}| \to 0}$$

and you're done. This obviously squeezes out the result $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$. Do this on exams! ∎

Theorem 2.11 *If $f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$, then f is continuous if and only if each of the component functions $f_i : \mathbb{R}^n \to \mathbb{R}$ is continuous.* ∎

The proof of this theorem is a simple consequence of the fact that

$$|x_i - a_i| \leq \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2} = |\mathbf{x} - \mathbf{a}| \leq \sqrt{n} \max_{1\leq i\leq n} |x_i - a_i|$$

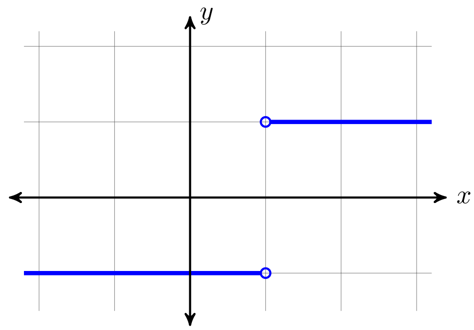
for all $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{x} = (x_1, \ldots, x_n)$ (Check this!) and then chasing the definition of continuity around. We omit the full tedious proof. The upshot of this very important theorem, however, is that we don't actually need to study the continuity of vector-valued functions to know what's going on. We need only study real-valued functions, for every vector-valued function is made up of real-valued component functions!

Theorem 2.12 *Sums, products, scalar multiples, quotients (so long as the denominator does not go to 0), and compositions of continuous functions are continuous.* ∎

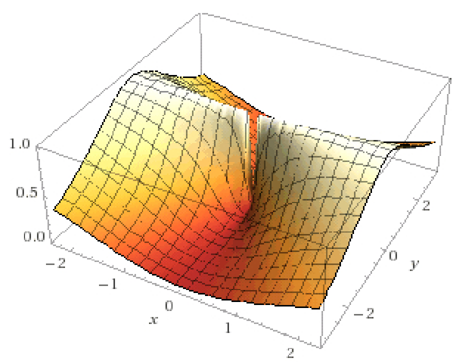
We also omit the proof of this theorem, as it's a lot of chasing ε's and δ's. I recommend taking the undergraduate analysis course or looking at one of the references below for full details.

Remark 2.13 *How does one show* **discontinuity***? There are three basic reasons a function f may be discontinuous.*

1. *First, the limit may not exist. A single variable example of this type of problem is encountered with the function* $f(x) = \frac{x-1}{|x-1|}$.



The limit from the left and the limit from the right do not agree at $x = 1$. *In two or more dimensions this problem manifests itself in even more varied ways, for the simple reason that a given point, say in* $\mathbb{R}^2$, *call it* $\mathbf{a} = (a, b)$, *may be approcated from more than two ways (not just from the left and the right). It can be approached along any path, even curved paths. A simple example of this is* $f(x, y) = \frac{x^2}{x^2+y^2}$. *Approaching* $(0, 0)$ *along the line* $x = 0$, *we see that* f *has a constant value of* 0, *while approaching* $(0, 0)$ *along the line* $y = x$ *we see that* f *has a constant value of* $1/2$. *These two are not the same* z*-values, so* f *cannot even have a limit at* $(0, 0)$, *even if we defined it to have some value there. You can see the rupture in the graph of* f:



2. *Second, the limit may exist, but the value of the function may not be equal to that of the limit. For example,*

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2+y^2}, & \text{if } (x, y) \neq (0, 0) \\ 1, & \text{if } (x, y) = (0, 0) \end{cases}$$

The limit as (x, y) *approaches* $(0, 0)$ *is* 0, *but the value of* f *at* $(0, 0)$ *is* 1. *This is analogous to the many Calc 1 examples, such as*

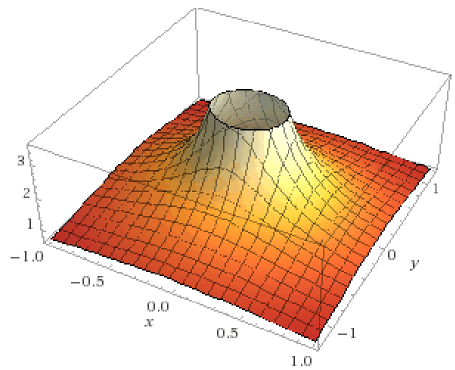
$$f(x) = \begin{cases} x^2, & \text{if } x \neq 2 \\ 1, & \text{if } x = 2 \end{cases}$$

This is called a **removable discontinuity**. *Let us prove that* $\lim_{(x,y)\to(0,0)} f(x, y) = 0$. *This is easy, for we have the inequality*

$$|f(x, y) - 0| = \left|\frac{x^2 y}{x^2 + y^2}\right| = \frac{x^2|y|}{x^2 + y^2} \leq |y| = \sqrt{y^2} \leq \sqrt{x^2 + y^2} = |(x, y) - (0, 0)|$$

on account of $\frac{x^2}{x^2+y^2} \leq 1$ *for all* (x, y) *in* $\mathbb{R}^2$*. By the above Remark 2.10 we are done, for* $|(x, y) - (0, 0)| \to 0$ *by assumption.*

3. *Third, the function is not defined there and/or escapes to infity or negative infinity. For example,* $f(x, y) = 1/\sqrt{x^2 + y^2}$ *as* (x, y) *goes to* $(0, 0)$*. This is exactly the two dimensional analog of the function* $f(x) = 1/|x|$*. You can see this by noting that* $\sqrt{x^2 + y^2} = |(x, y)|$*, so the function is* $f(\mathbf{x}) = 1/|\mathbf{x}|$*, in fact.* ■



2.3 Further Examples

Example 2.14 $f(x,y) = \dfrac{e^{\sin x}}{\cos y}$ *is continuous on the square* $[-\frac{\pi}{2}, \frac{\pi}{2}] \times [0, \frac{\pi}{4}]$*, because* $\cos y$ *is nonzero and continuous on* $[0, \frac{\pi}{4}]$*, and it's reciprocal is nonzero there, and also* $\sin x$ *and therefore* $e^{\sin x}$ *is continous, so their quotient is continuous.* ■

Example 2.15 $f(x,y) = \dfrac{x+y}{x-y}$ *is not continuous at* $(0,0)$*, for if we approach* $(0,0)$ *along the line* $x = -y$*, we get* $f(x,y) = f(-y,y) = \frac{-2y}{-2y} = -1$*, while if we approach* $(0,0)$ *along the line* $x = 2y$ *we get* $f(x,y) = f(2y,y) = \frac{3y}{y} = 3$. ■

Example 2.16 $f(x,y) = \dfrac{x^2 - y^2}{x^2 + y^2}$ *is not continuous at* $(0,0)$*, for if we approach* $(0,0)$ *along the line* $x = y$ *we get that* $f(x,y) = f(x,x) = 0$*, while if we approach* $(0,0)$ *along the line* $x = 2y$ *we get that* $f(x,y) = f(2y,y) = \frac{3y^2}{5y^2} = \frac{3}{5}$. ■

Example 2.17 $f(x,y) = \dfrac{xy}{|xy|}$ *is not continuous at* $(0,0)$*. Taking* $x, y > 0$*, then* $f(x,y) = 1$*, while taking* $x > 0$ *and* $y < 0$*, say, we have* $f(x,y) = -1$. ■

Example 2.18 $f(x,y) = \dfrac{x^2}{x^2 + y}$ *is not continuous at* $(0,0)$*, for look at the curves* $y = 0$*,* $y = x^2$ *and* $y = 2x^2$*.* f *has values* 1*,* $1/2$ *and* $1/3$*, respectively, on those curves as it approaches* $(0,0)$. ■

Example 2.19 *The function* $f(x,y) = 3x - 2y + 5$ *is continuous everywhere, for it is a polynomial, indeed it is a sum of scalar multiples of simple monomials. But we could also prove it's continuity by use of the triangle inequality*

$$\begin{aligned}
|f(x,y) - f(a,b)| &= |(3x - 2y + 5) - (3a - 2b + 5)| \\
&= |3(x-a) - 2(y-b)| \\
&\le |3(x-a)| + |-2(y-b)| \\
&= 3|x-a| + 2|y-b| \\
&= 3\sqrt{(x-a)^2} + 2\sqrt{(y-b)^2} \\
&\le 3\sqrt{(x-a)^2 + (y-b)^2} + 2\sqrt{(x-a)^2 + (y-b)^2} \\
&= (3+2)\sqrt{(x-a)^2 + (y-b)^2} \\
&= 5|(x,y) - (a,b)| \to 0
\end{aligned}$$

Example 2.20 *Let* $f(x,y) = \begin{cases} \dfrac{x^3y + xy^3}{2x^2 + 2y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$*. Then* f *is continuous everywhere, including* $(0,0)$*, for so long as* $(x,y) \neq (0,0)$ *we can simplify*

$$f(x,y) = \frac{x^3y + xy^3}{2x^2 + 2y^2} = \frac{xy(x^2 + y^2)}{x^2 + y^2} = xy$$

and the limit $\lim_{(x,y)\to(0,0)} xy = 0$. ■

Example 2.21 *Let* $f(x,y) = \begin{cases} \dfrac{\sin(2\sqrt{x^2+y^2})}{3\sqrt{x^2+y^2}}, & (x,y) \neq (0,0) \\ \frac{2}{3}, & (x,y) = (0,0) \end{cases}$. *Then* f *is continuous everywhere, including* $(0,0)$*, for if we use polar coordinates we get*

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} \frac{\sin(2\sqrt{x^2+y^2})}{3\sqrt{x^2+y^2}} = \lim_{r\to 0} \frac{\sin 2r}{3r} = \frac{2}{3} \lim_{r\to 0} \frac{\sin 2r}{2r} = \frac{2}{3} = f(0,0) \quad \blacksquare$$

References

[1] Burkill, J. C., and H. Burkill, *A Second Course in Mathematical Analysis*, Cambridge University Press, 1970

[2] Duistermaat, J.J., Kolk, J.A.C., *Multidimensional Real Analysis I: Differentiation*, Cambridge, 2004

[3] Duistermaat, J.J., Kolk, J.A.C., *Multidimensional Real Analysis 2: Integration*, Cambridge, 2004

[4] Friedberg, Stephen H., Arnold J. Insel, Lawrence E. Spence, *Linear Algebra*, 4th Ed., Prentice Hall, 2002

[5] Grove, Larry C., *Classical Groups and Geometric Algebra*, American Mathematical Society, 2002

[6] Lang, Serge, *Undergraduate Analysis*, 2nd Ed., Springer, 1997

[7] Rudin, Walter, *Principles of Mathematical Analysis*, 3rd Ed., McGraw-Hill, 1976