Lana's Better CALC II Lecture Notes

Lana Mantegazza

February 24, 2026

Preface

This is my attempt at making a more comprehensable set of lecture notes for the CALC II module. Due to its overall better legibility and structure, I've based the style of these lecture notes on that used in the LAG II lecture notes. Special thanks to me for spending a full ass day painstakingly reconstructing the LaTeX preamble used in the LAG II lecture notes, I hope I've done a good enough job and that these lecture notes are at least slightly better than the ones provided by our module. Basically, after I say anything just imagine it says "*From what I've been able to gather*" before it. Also this is my first time using LaTeX so hope it all looks good and up to code.

– Lana Mantegazza M.D.[1] Ph.D[1]

[1]probably

Contents

Introduction and Overview

0.1 The Purpose of These Notes

0.2 A Brief Review of Calculus 1

0.3 A Brief Review of Linear Algebra and Geometry

As this module is focused on **vector** calculus, it is important to have a solid foundation in linear algebra and geometry. Let us briefly review some of the key concepts from Linear Algebra and Geometry I.

0.3.1 Vectors

Definition 0.3.1. Let $S = \{v_1, v_2, \cdots, v_n \mid v_i \in \mathbb{R}, \forall i \leq n \in \mathbb{N}\}$ be a set of n real numbers. Then an n-dimensional **vector** $\mathbf{v} \in \mathbb{R}^n$ is the ordered tuple of the elements of S

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \tag{0.3.1}$$

sometimes written as

$$\mathbf{v} = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}^T \tag{0.3.2}$$

Remark 0.3.2. For the purposes of this document, if a vector $\mathbf{v}$ in $\mathbb{R}^n$ is declared without being explicitly defined, you may assume that its entries are

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Definition 0.3.3. Let $\mathbf{v}$ be a vector. Then the **scalar multiplication** of $\mathbf{v}$ by a real number $c \in \mathbb{R}$ is the vector

$$c\mathbf{v} = \begin{pmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{pmatrix} \tag{0.3.3}$$

Definition 0.3.4. Let $\mathbf{v}$ be a vector in $\mathbb{R}^n$, Then the **magnitude** of $\mathbf{v}$ is the real number

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \tag{0.3.4}$$

Definition 0.3.5. A **unit vector** is a vector of magnitude 1.

Definition 0.3.6. Let $\mathbf{v} \in \mathbb{R}^n$ be a vector. Then to **normalize v** is to take the unit vector

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|} \tag{0.3.5}$$

Remark 0.3.7. When defining vectors in $\mathbb{R}^3$, it is common to denote the first, second, and third entries in a vector as coefficients of the unit vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ respectively, where

$$\hat{\mathbf{i}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \hat{\mathbf{j}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \hat{\mathbf{k}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{0.3.6}$$

Definition 0.3.8. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ be two vectors. Then the **dot product** of $\mathbf{v}$ and $\mathbf{w}$ is the real number

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n \tag{0.3.7}$$

Definition 0.3.9. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ be two vectors in $\mathbb{R}^3$. Then the **cross product** of $\mathbf{v}$ and $\mathbf{w}$ is the vector

$$\mathbf{v} \times \mathbf{w} = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} = \det \begin{pmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} \tag{0.3.8}$$

A brief word on the Hadamard product of two vectors. The Hadamard product of two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$ is what we all thought multiplying vectors would look like before we learned about dot products and cross products. You take each entry of $\mathbf{v}$ and multiply it with the corresponding entry of $\mathbf{w}$, and you have the corresponding entry in the result.

Definition 0.3.10. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ be two vectors. Then the **Hadamard product** of $\mathbf{v}$ and $\mathbf{w}$ is the vector

$$\mathbf{v} \odot \mathbf{w} = \begin{pmatrix} v_1 w_1 \\ v_2 w_2 \\ \vdots \\ v_n w_n \end{pmatrix} \tag{0.3.9}$$

0.3.2 Matrices

Matrices are a way of representing linear transformations between vector spaces, and are useful for pretty much everything in vector calculus.

Definition 0.3.11. Let $m, n \in \mathbb{N}$ be natural numbers. Then an m by n **matrix** is a rectangular array of real numbers with m rows and n columns

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{0.3.10}$$

Definition 0.3.12. Let A be an m by n matrix. Then the **transpose** of A is the n by m matrix

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix} \tag{0.3.11}$$

Remark 0.3.13. If A is an m by n matrix and $m = n$, then A is a square matrix.

Definition 0.3.14. Let A be an m by n matrix and let B be an n by p matrix. Then the **matrix product** of A and B is the m by p matrix

$$\begin{aligned} C = AB &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix} \\ &= \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{pmatrix} \end{aligned} \tag{0.3.12}$$

such that

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \quad \forall i \leq m, j \leq p \tag{0.3.13}$$

0.3.3 Geometry

A quick reminder of the definition of some basic geometric properties of objects in $\mathbb{R}^n$.

Definition 0.3.15. Let $\mathbf{v}$ and $\mathbf{w}$ be two vectors in $\mathbb{R}^n$. Then the **angle** θ between $\mathbf{v}$ and $\mathbf{w}$ is such that

$$\cos(\theta) = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} \tag{0.3.14}$$

If $\mathbf{v}$ and $\mathbf{w}$ are vectors in $\mathbb{R}^3$, then the **angle** θ between $\mathbf{v}$ and $\mathbf{w}$ is also given by

$$\sin(\theta) = \frac{\|\mathbf{v} \times \mathbf{w}\|}{\|\mathbf{v}\|\|\mathbf{w}\|} \tag{0.3.15}$$

Definition 0.3.16. Let $\mathbf{v}$ and $\mathbf{w}$ be two vectors in $\mathbb{R}^3$. Then $\mathbf{v}$ is **parallel** to $\mathbf{w}$ if and only if

$$\mathbf{v} \times \mathbf{w} = \mathbf{0} \tag{0.3.16}$$

and $\mathbf{v}$ is **normal** to $\mathbf{w}$ if and only if

$$\mathbf{v} \cdot \mathbf{w} = 0 \tag{0.3.17}$$

Definition 0.3.17. Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be three vectors in $\mathbb{R}^3$. Then $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are **coplanar** if and only if

$$\mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = 0 \tag{0.3.18}$$

Remark 0.3.18. Finally, it is important to have an informal idea of what tangents and normals are. We will define them more formally in Chapters 1 and 2, but for now, refer to the following informal definitions.

Definition 0.3.19. Informal Definition of a Tangent

Let C be a curve in $\mathbb{R}^n$, and let p be a point on C. Then the **tangent** to C at p is the line which just touches C at p, and is parallel to the direction in which C at p.

Definition 0.3.20. Informal Definition of a Normal

Let S be a surface in $\mathbb{R}^n$, and let p be a point on S. Then the **normal** to S at p is the line which just touches S at p, and is perpendicular to the tangent plane to S at p.

0.4 The Four Types of Functions

In Calculus 1 and Calculus 2, our primary focus has been and will continue to be studying functions. The aim of this module is to expand the domain of what we learned previously in Calculus I to higher dimensions, and there are four distinct types of functions which we will consider to achieve this aim.

$$\begin{array}{r|l} \mathbf{r} : \mathbb{R}^1 \longrightarrow \mathbb{R}^n & \text{(paths in } \mathbb{R}^n\text{)} \\ f : \mathbb{R}^m \longrightarrow \mathbb{R}^1 & \text{(scalar value functions on } \mathbb{R}^m\text{)} \\ \mathbf{v} : \mathbb{R}^n \longrightarrow \mathbb{R}^n & \text{(vector fields on } \mathbb{R}^n\text{)} \\ T : \mathbb{R}^m \longrightarrow \mathbb{R}^n & \text{(vector functions from } \mathbb{R}^m \text{ to } \mathbb{R}^n\text{)} \end{array}$$

For each of these functions, we will explore how they are defined, how we can expand our definitions of derivatives and integrals to encompass them, and how these both relate to higher dimensional geometry. Moreover, we will see how calculus, with relation to these functions, gives rise to methods for computing useful geometric properties of objects in higher dimensional space.

Chapter 1

Paths and Parametric Equations

Let us assume that $n \in \mathbb{N}$ is a natural number such that $n > 1$. This chapter will cover functions

$$\mathbf{r} : \mathbb{R} \longrightarrow \mathbb{R}^n \tag{1.0.1}$$

mapping from $\mathbb{R}^1$ to $\mathbb{R}^n$, usually referred to as paths or, equivalently, as parametric equations

$$t \longmapsto (f(t), g(t)) \tag{1.0.2}$$

1.1 Paths and Curves

There is an important distinction to be made between paths and curves. While paths in $\mathbb{R}^n$ are functions, curves in $\mathbb{R}^n$ are instead geometric objects in n dimensional space.

Definition 1.1.1. Let $\mathbf{r} : \mathbb{R} \longrightarrow \mathbb{R}^n$ be a function. Then $\mathbf{r}$ is a **path** on $\mathbb{R}^n$

Definition 1.1.2. Let $C \subseteq \mathbb{R}^n$ be a subset of points in $\mathbb{R}^n$. If there exists some path $\mathbf{r} : \mathbb{R} \longrightarrow C \subseteq \mathbb{R}^n$ such that $\mathbf{r}$ is continuous, then C is a **curve**.

Chapter 2

Scalar Value Functions

Let us assume that $m \in \mathbb{N}$ is a natural number such that $m > 1$. This chapter will cover functions

$$f : \mathbb{R}^m \longrightarrow \mathbb{R} \tag{2.0.1}$$

mapping from $\mathbb{R}^m$ to $\mathbb{R}$, referred to as scalar value functions on $\mathbb{R}^m$, which are often used to define surfaces in $\mathbb{R}^{m+1}$. We will mostly be working with scalar value functions on $\mathbb{R}^2$.

Note: Most definitions will write functions as mapping $\mathbf{v} \longmapsto f(\mathbf{v})$. Keep in mind that writing $f(\mathbf{v})$ is equivalent to writing $f(x, y)$ when $m = 2$, and that the use of $\mathbf{v}$ is to allow definitions to encompass higher dimensions.

2.1 Scalar Value Functions and Surfaces

Much like paths and curves, scalar value functions differ from surfaces. Scalar value functions are functions mapping from $\mathbb{R}^m$ to $\mathbb{R}$, while surfaces are geometric objects in $m+1$ dimensional space.

Definition 2.1.1. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a function. Then f is a **scalar value function** on $\mathbb{R}^m$.

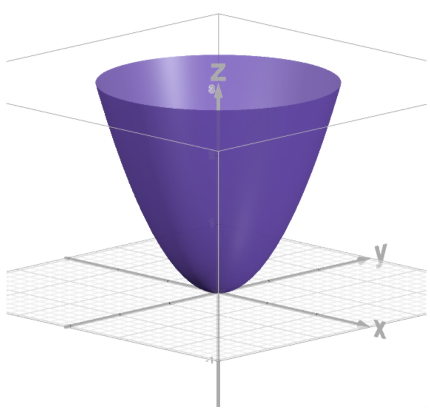
Definition 2.1.2. Let $S \subseteq \mathbb{R}^{m+1}$ be a subset of points in $\mathbb{R}^{m+1}$. Then S is the **surface** described by the scalar value function $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ if and only if

$$S = \left\{ \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \\ f(\mathbf{v}) \end{pmatrix} \in \mathbb{R}^{m+1} \right\} \tag{2.1.1}$$

Example 2.1.3. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 + y^2$ be a scalar value function. Then the surface S described by f is the set of points

$$S = \left\{ \begin{pmatrix} x \\ y \\ x^2 + y^2 \end{pmatrix} \in \mathbb{R}^3 \right\}$$

This specific surface is called a paraboloid, seen in the figure below.



Example 2.1.4. A Geometric Interpretation of Scalar Value Functions

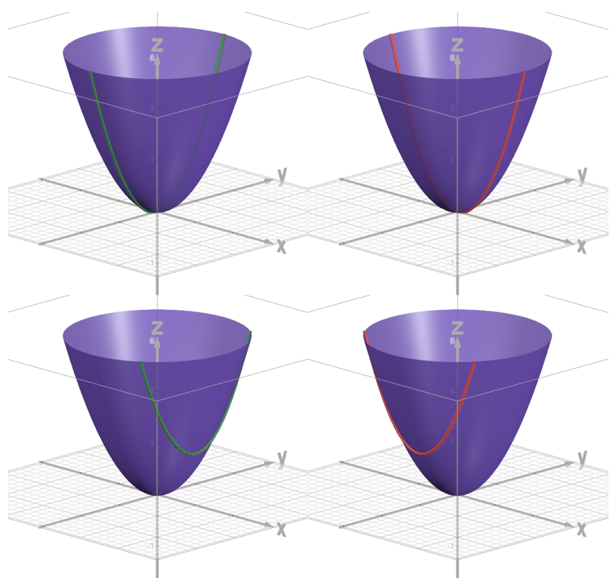
When we plot this surface in 3D space, we see that it is what is called a paraboloid, seen in 2.1.3, but it may not be immediately clear why this is the case. However, if we consider how this surface looks when viewed from different prespectives, it is more apparent as to why this is the case.

Consider the scalar value function

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 + y^2$$

When we look at the surface described by f from either side, we can see that each slice along the x or y axis is just a different parabola. For example,

- at $x = 0$, we have the parabola $z = y^2$,
- at $y = 0$, we have the parabola $z = x^2$,
- at $x = 1$, we have the parabola $z = y^2 + 1$,
- at $y = 1$, we have the parabola $z = x^2 + 1$,



This geometric interpretation of scalar value functions will be integral[1] to our understanding the geometric properties of partial derivatives, covered later in this chapter.

[1]Pun not intended.

2.2 Algebraic and Analytic Surfaces

An equivalent way of defining a surface in $\mathbb{R}^{m+1}$ is to define it as the set of points in $\mathbb{R}^{m+1}$ which satisfy an algebraic equation with $m+1$ variables. We call these **algebraic surfaces.**

Definition 2.2.1. Let $S \subseteq \mathbb{R}^{m+1}$ be a surface in $\mathbb{R}^{m+1}$, and let $g : \mathbb{R}^{m+1} \longrightarrow \mathbb{R}$ be a function. Then S is the **algebraic surface** described by g if and only if

$$S = \left\{ \mathbf{v} \in \mathbb{R}^{m+1} : g(\mathbf{v}) = 0 \right\} \tag{2.2.1}$$

To make this clearer, let us consider the a case in $\mathbb{R}^3$.

Example 2.2.2. Let $x^2 - z = -y^2$ be an algebraic equation in $\mathbb{R}^3$. We can rearrange this equation

$$\begin{aligned} x^2 - z &= -y^2 \\ x^2 + y^2 &= z \end{aligned} \tag{2.2.2}$$

and see that this algebraic equation is equivalent to the scalar value function

$$f(x, y) = x^2 + y^2$$

and thus the algebraic surface described by the equation $x^2 + y^2 - z = 0$ is the same as that of $f(x, y) = x^2 + y^2$, which is the paraboloid seen in 2.1.3.

Remark 2.2.3. In the case where our algebraic surface is defined by an algebraic equation of the form $g(x, y, z) = c$, for some constant $c \in \mathbb{R}$, we call the algebraic surface described by g an **analytic surface**.

2.3 Level Sets and Level Curves

When working with scalar value functions, it is often useful to consider what is called a level set, or, in the context of scalar value functions on $\mathbb{R}^2$, a level curve.

Definition 2.3.1. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a scalar value function on $\mathbb{R}^m$, and let $c \in \mathbb{R}$ be a real number. Then the **level set of f at c** is the set of points in $\mathbb{R}^m$ which are mapped to c by f

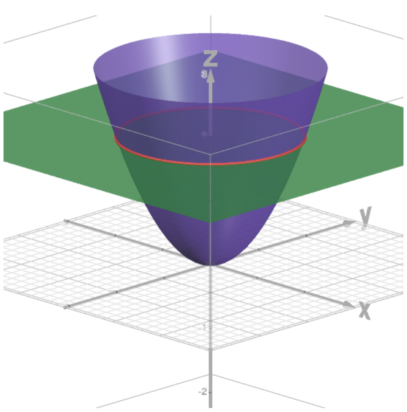
$$L_c = \{\mathbf{v} \in \mathbb{R}^m : f(\mathbf{v}) = c\} \tag{2.3.1}$$

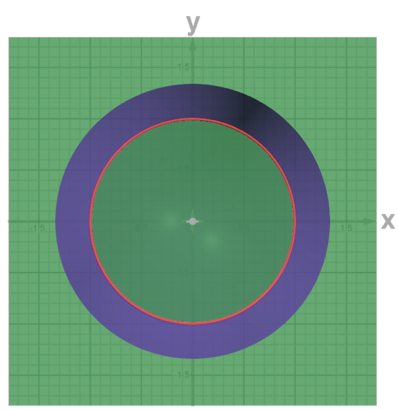
Remark 2.3.2. A level curve can be thought of as the *cross-section* of a surface with a plane. In example 2.3.3, we will visualize the level curve of f at c as such, seeing how the level curve of f at c is the intersection of the surface described by f and the plane $g(x, y) = c$.

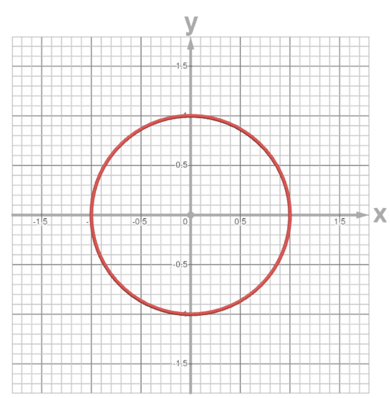
Example 2.3.3. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 + y^2$ be a scalar value function, and let $g(x, y) = c = 1$. Then the **level curve of f at c** is the set of points in $\mathbb{R}^2$

$$L_1 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 + y^2 = 1 \right\}$$

When thinking about the points $(x, y) \in \mathbb{R}^2$ such that $x^2 + y^2 = 1$, it is useful to recall the 2D plot of the algebraic curve $x^2 + y^2 = 1$. This is indeed the level curve of f at c for $f(x, y) = x^2 + y^2$. The figures below show the paraboloid S described by f in purple, the plane $g(x, y) = c$ in green, and the **level curve L_1 of f at c** in **red**.







2.4 Partial Derivatives

To make our definitions more tractible to read, let us denote the unit vector in the i^{th} direction of our vector space $\mathbb{R}^m$ as

$$\hat{\mathbf{v}}_i = \begin{pmatrix} 0 & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}^T \tag{2.4.1}$$

where the i^{th} entry is 1, and all other entries are 0.

Definition 2.4.1. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a scalar value function on $\mathbb{R}^m$. Then the **partial derivative of f with respect to v_i** is

$$\frac{\partial f}{\partial v_i}(\mathbf{v}) = \lim_{h \to 0} \frac{f(\mathbf{v} + h\hat{\mathbf{v}}_i) - f(\mathbf{v})}{h} \tag{2.4.2}$$

While this symbol-soup may look weird as hell, it is a lot easier to grasp when we consider the partial derivatives of a scalar value function on $\mathbb{R}^2$. As the majority of the content for this module deals with scalar value functions in $\mathbb{R}^2$, it will be more useful to refer to Example 2.4.2 for the first-principles definitions of partial derivatives.

Example 2.4.2. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 + y^2$ be a scalar value function on $\mathbb{R}^2$. Then the **partial derivative** of f with respect to x is

$$f_x(x, y) = \frac{\partial f}{\partial x}(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h} \tag{2.4.3}$$

and the partial derivative of f with respect to y is

$$f_y(x, y) = \frac{\partial f}{\partial y}(x, y) = \lim_{h \to 0} \frac{f(x, y + h) - f(x, y)}{h} \tag{2.4.4}$$

Despite the complicated definition of partial derivatives, computing them is done rather easily. The procedure outlined below is for scalar value functions on $\mathbb{R}^2$, but can be easily generalized to higher dimensions.

Procedure 2.4.3. Computing Partial Derivatives

Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto f(x, y)$ be a scalar value function on $\mathbb{R}^2$ and let $c \in \mathbb{R}$ be some constant. Then

$$\frac{\partial f}{\partial x} = \frac{d}{dx} f(x, c) \tag{2.4.5}$$

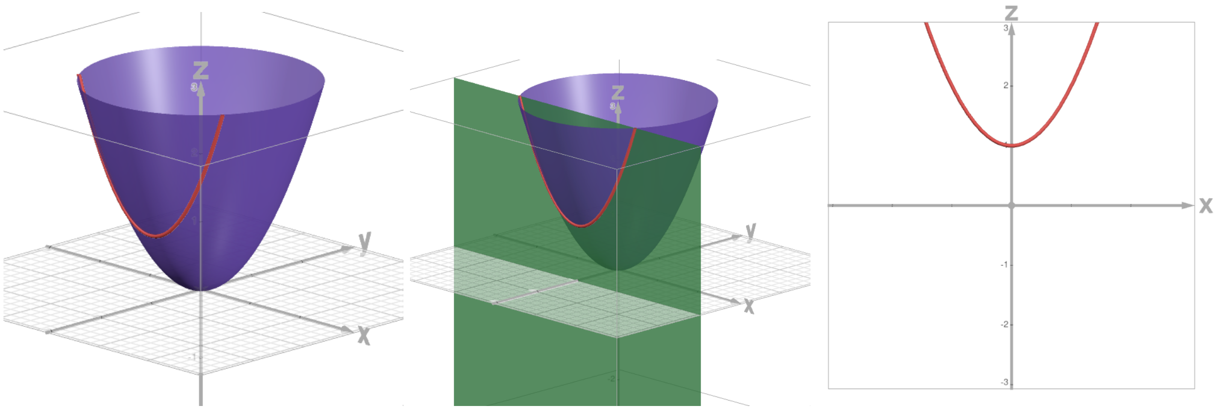
Example 2.4.4. A Geometric Interpretation of Partial Derivatives

It is useful to have a geometric understanding of what the partial derivative of a scalar value function represents.

Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 + y^2$ be a scalar value function. We know this function generates a paraboloid. Recall that, when viewing the paraboloid from the side, we can see that each slice along the x or y axis is a parabola in $\mathbb{R}^2$ (see 2.1.4).

Consider the slice of the paraboloid along the x axis at $y = 1$

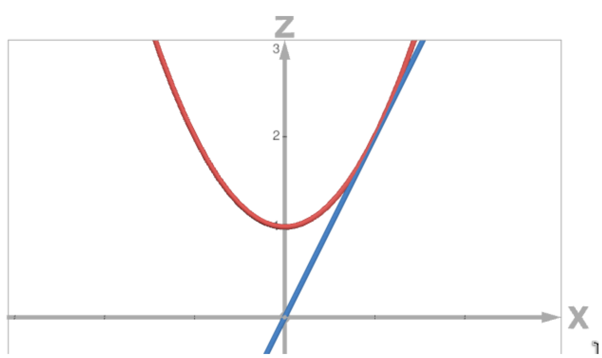
$$f(x, 1) = x^2 + 1 \tag{2.4.6}$$



If we now consider the partial derivative of f with respect to x

$$\frac{\partial}{\partial x} f(x, y) = 2x$$

we can see that plotting it reveals that it is in fact just the derivative $\frac{d}{dx}$ of the aforementioned parabola (2.4.6).



2.5 The Gradient and The Nabla Operator

When considering scalar value functions from $\mathbb{R}^m$ to $\mathbb{R}$, it is useful to be able to denote the vector in $\mathbb{R}^m$ containing the partial derivatives of f with respect to each of its inputs.

Definition 2.5.1. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R} : \mathbf{v} \longmapsto f(\mathbf{v})$ be a scalar value function on $\mathbb{R}^m$. Then the **gradient** of f is the vector function

$$\nabla f : \mathbb{R}^m \longrightarrow \mathbb{R}^m : \mathbf{v} \longmapsto \begin{pmatrix} \frac{\partial}{\partial v_1} f(\mathbf{v}) \\ \frac{\partial}{\partial v_2} f(\mathbf{v}) \\ \vdots \\ \frac{\partial}{\partial v_m} f(\mathbf{v}) \end{pmatrix} \tag{2.5.1}$$

Remark 2.5.2. The upside-down triangle symbol (∇) is called the **nabla operator**, and is not exclusively used to denote the gradient of a scalar value function. In what is – for some reason – referred to as an "abuse[2] of notation" in some literature surrounding vector calculus, the nabla operator is given its own definition which is very helpful to have for defining other types of derivatives that we will explore later in this module.

Definition 2.5.3. The **nabla operator** for the vector space $V \subseteq \mathbb{R}^m$ with bases $\{\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \cdots, \hat{\mathbf{v}}_m\}$ is defined to be the vector

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial v_1} \\ \frac{\partial}{\partial v_2} \\ \vdots \\ \frac{\partial}{\partial v_m} \end{pmatrix} \tag{2.5.2}$$

Remark 2.5.4. With this definition of the nabla operator, we can now also understand the gradient of a scalar value function as the *Hadamard product* (see 0.3.10) of the nabla operator and the scalar value function itself.

[2]No notations were harmed in the making of these lecture notes

Example 2.5.5. A Geometric Interpretation of the Gradient

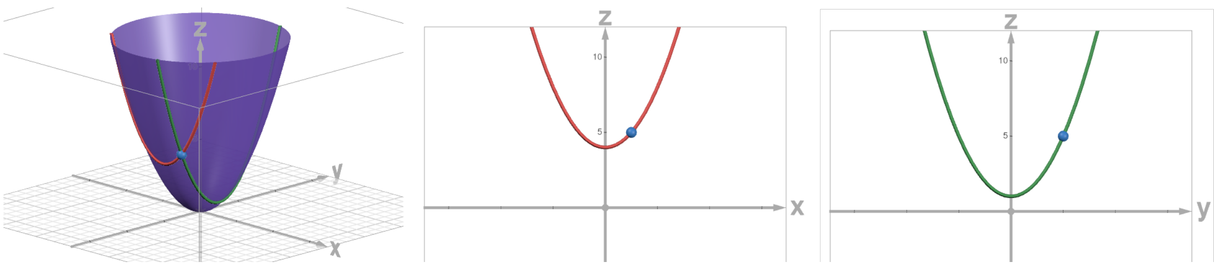
Let us briefly reflect on how one might visualize the gradient of a scalar value function on $\mathbb{R}^2$ through the use of a 3D plot.

Let us once again consider the scalar value function $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 + y^2$. We know that its **gradient** is given by

$$\nabla f(x, y) = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

and that the surface described by this scalar value function is a **paraboloid** (see 2.1.4), but how do these two things relate to each other?

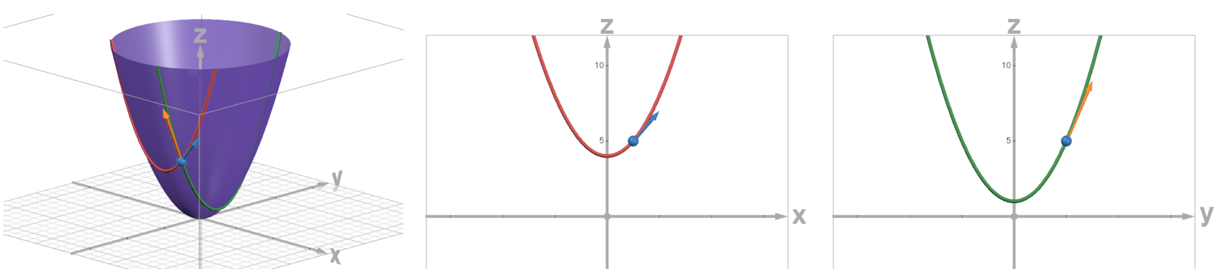
Let us now consider the point $(1, 2) \in \mathbb{R}^2$, mapping to the point $(1, 2, 5) \in \mathbb{R}^3$ on our paraboloid.



Computing the gradient of f gives us

$$\nabla f(1, 2) = \begin{pmatrix} f_x(1, 2) \\ f_y(1, 2) \end{pmatrix} = \begin{pmatrix} 2(1) \\ 2(2) \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

The key point here is to realize that the partial derivatives of f, f_x and f_y, are **exactly** the slopes of f at the point $(1, 5)$ when viewed from only the x and z axes, and at $(2, 5)$ when viewed from only the y and z axes, and we can see that by drawing the tangent lines with these slopes onto their respective curves.



2.6 Directional Derivatives

While partial derivatives are useful for getting the rate of change of some scalar value function $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ with respect to one of the axes x or y, how might one go about finding the derivative of a scalar value function in any arbitrary direction?

Definition 2.6.1. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R} : \mathbf{v} \longmapsto f(\mathbf{v})$ be a scalar value function on $\mathbb{R}^m$, and let $\mathbf{u} \in \mathbb{R}^m$ be a unit vector. Then the **directional derivative of f in the direction of u** at some point $\mathbf{v} \in \mathbb{R}^{m+1}$ is the rate of change of f in the direction of $\mathbf{u}$ at the point $\mathbf{v}$.

Theorem 2.6.2. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a scalar value function on $\mathbb{R}^m$, and let $\mathbf{u} \in \mathbb{R}^m$ be a unit vector. Then the **directional derivative of f in the direction of u** is

$$\nabla_{\mathbf{u}} f = \nabla f \cdot \mathbf{u} \tag{2.6.1}$$

Proof of Theorem 2.6.2 for scalar value functions on $\mathbb{R}^2$.
We must prove that $\nabla_{\mathbf{u}} f(\mathbf{v}) = \nabla f(\mathbf{v}) \cdot \mathbf{u}$ is the rate of change of f in the direction of $\mathbf{u}$ at the point $\mathbf{p} \in \mathbb{R}^3$.

Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : \mathbf{v} \longmapsto f(\mathbf{v})$ be a scalar value function on $\mathbb{R}^2$, and let $\mathbf{u} \in \mathbb{R}^m$ be a unit vector. Furthermore, let $\mathbf{p} = (p_x, p_y), \mathbf{u} = (u_x, u_y), \hat{\mathbf{i}} = (1, 0),$ and $\theta = \angle(\mathbf{u}, \hat{\mathbf{i}})$

1. Recall the definition of a unit vector from 0.3.5, $\mathbf{u}$ is a unit vector $\implies \|\mathbf{u}\| = 1$. Moreover, recall that, by the definition of the dot product 0.3.15,

$$\begin{aligned}
&\|\mathbf{u}\| \|\hat{\mathbf{i}}\| \cos\theta = \mathbf{u} \cdot \hat{\mathbf{i}} \\
&\quad \implies \cos\theta = u_x \\
&\quad \implies \sin\theta = u_y = \sqrt{1 - u_x^2} \\
&\quad \implies \exists T = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ s.t. } T\mathbf{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \iff T = \begin{pmatrix} u_x & -u_y \\ u_y & u_x \end{pmatrix}
\end{aligned} \tag{2.6.2}$$

2. Let us transform our function such that our point $\mathbf{p}$ is at the origin, and our direction $\mathbf{u}$ lies along the basis vector $\hat{\mathbf{i}}$. We can do this by first translating our vector space by $-\mathbf{p}$, and then rotating our vector space to put $\mathbf{u}$ along a basis.

   Let $g : \mathbb{R}^m \longrightarrow \mathbb{R}$ be our transformed function

$$\begin{aligned}
&g(\mathbf{v}) = f(T\mathbf{v} + \mathbf{p}) \\
\implies &g(\mathbf{0}) = f(\mathbf{p}), \text{ and} \\
&g(\mathbf{u}) = f(\hat{\mathbf{i}}).
\end{aligned} \tag{2.6.3}$$

3. Now, let us consider the directional derivative of $g(\mathbf{v})$ in the direction $T\mathbf{u}$

   As $T\mathbf{u} = \hat{\mathbf{i}}$, the directional derivative of $g(\mathbf{v})$ in the direction $\mathbf{u}$ must be equal to the partial derivative of $g(x, y)$ with respect to x at $\mathbf{0}$

$$\begin{aligned}
\nabla_{\mathbf{u}} f(\mathbf{v}) &= \nabla_{\hat{\mathbf{i}}} g(\mathbf{0}) \\
&= \frac{\partial g}{\partial x}(\mathbf{0}) \\
&= \lim_{h \to 0} \frac{g(h,0) - g(0,0)}{h} \\
&= \lim_{h \to 0} \frac{f\left(\begin{pmatrix} u_x & u_y \\ u_y & -u_x \end{pmatrix}\begin{pmatrix} h \\ 0 \end{pmatrix} + \mathbf{p}\right) - f(\mathbf{p})}{h} \\
&= \lim_{h \to 0} \frac{f(h\mathbf{u} + \mathbf{p}) - f(\mathbf{p})}{h} \\
&= \lim_{h \to 0} \left(\frac{f(hu_x + p_x, p_y) - f(p_x, p_y)}{h}\right) + \lim_{h \to 0} \left(\frac{f(p_x, hu_y + p_y) - f(p_x, p_y)}{h}\right) \\
&= u_x \frac{\partial f}{\partial x}(\mathbf{p}) + u_y \frac{\partial f}{\partial y}(\mathbf{p}) \\
&= \mathbf{u} \cdot \nabla f(\mathbf{p})
\end{aligned} \tag{2.6.4}$$

□

Now, equipped with these tools, we can explore how we can use calculus to compute some useful geometric properties of the surfaces described by these scalar value functions.

2.7 Tangent Vectors to Surfaces

Much like how paths have a normalized tangent vector at each point, and a plane of possible normal vectors at each point, surfaces have a single normalized normal vector at each point, and a plane of possible tangent vectors at each point. First, let us farmiliarize ourselves with the definition of tangent vectors to surfaces.

Definition 2.7.1. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a scalar value function on $\mathbb{R}^2$ describing the surface $S \subseteq \mathbb{R}^3$, and let $\mathbf{p} \in \mathbb{R}^2$ be some point on S. If

$$\exists \mathbf{u} = \begin{pmatrix} u_x \\ u_y \end{pmatrix} \in \mathbb{R}^2 \text{ such that } \mathbf{v} = \begin{pmatrix} u_x \\ u_y \\ \nabla_{\mathbf{u}} f(\mathbf{p}) \end{pmatrix}, \tag{2.7.1}$$

then $\mathbf{v}$ is **tangent** to the surface S at the point $\mathbf{p}$.

Definition 2.7.2. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a scalar value function on $\mathbb{R}^2$ describing the surface $S \subseteq \mathbb{R}^3$, and let $\mathbf{p} \in \mathbb{R}^2$ be some point on S. Then the set of all points $\mathbf{v} \in \mathbb{R}^3$ tangent to the surface S at the point $\mathbf{p}$ is the **tangent plane** to S at $\mathbf{p}$

$$P = \left\{ \mathbf{v} \in \mathbb{R}^3 : \mathbf{v} = \begin{pmatrix} u_x \\ u_y \\ \nabla_{\mathbf{u}} f(\mathbf{p}) \end{pmatrix}, \forall \mathbf{u} \in \mathbb{R}^2 \right\} \tag{2.7.2}$$

Remark 2.7.3. Here, we are defining tangent vectors to a surface at a point in terms of the directional derivative of the scalar value function describing that surface. While this may seem like an overly complicated it also gives insight into a useful corollary of 2.6.2.

Corollary 2.7.4. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a scalar value function on $\mathbb{R}^m$, and let $S \subseteq \mathbb{R}^{m+1}$ be the surface described by f. Then the vector

$$\mathbf{v} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \\ \nabla_{\mathbf{u}} f(\mathbf{p}) \end{pmatrix} \tag{2.7.3}$$

is tangent to the surface S at the point $\mathbf{p}$ for all $\mathbf{u} \in \mathbb{R}^m$.

From this we can derive yet another corollary concerning the partial derivatives of a scalar value function themselves. We will write this one out for the case of scalar value functions on $\mathbb{R}^2$, but it can be generalized to higher dimensions.

Corollary 2.7.5. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a scalar value function on $\mathbb{R}^2$, and let $S \subseteq \mathbb{R}^3$ be the surface described by f. Then the vectors

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ f_x \end{pmatrix} \text{ and } \mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ f_y \end{pmatrix} \tag{2.7.4}$$

are tangent to the surface S at the point $\mathbf{p}$.

2.8 Normal Vectors to Surfaces

Now that we have defined tangent vectors to surfaces, we can define normal vectors to surfaces as well.

Definition 2.8.1. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a scalar value function on $\mathbb{R}^m$, and let $S \subseteq \mathbb{R}^{m+1}$ be the surface described by f. If $P \subseteq \mathbb{R}^{m+1}$ is the tangent plane to S at some point $\mathbf{p} \in \mathbb{R}^m$, then a vector $\mathbf{n} \in \mathbb{R}^{m+1}$ is a **normal vector** to S at $\mathbf{p}$ if and only if

$$\mathbf{n} \cdot \mathbf{v} = 0, \forall \mathbf{v} \in P \tag{2.8.1}$$

i.e. $\mathbf{n}$ is orthogonal to all tangent vectors to S at $\mathbf{p}$.

Theorem 2.8.2. Let $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ be a scalar value function on $\mathbb{R}^m$, and let $S \subseteq \mathbb{R}^{m+1}$ be the surface described by f. Then the **normal vector** to S at some point $\mathbf{v} \in \mathbb{R}^m$ is the vector

$$\mathbf{n} = \begin{pmatrix} \uparrow \\ \nabla f \\ \downarrow \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial v_1} f(\mathbf{v}) \\ \frac{\partial}{\partial v_2} f(\mathbf{v}) \\ \vdots \\ \frac{\partial}{\partial v_m} f(\mathbf{v}) \\ -1 \end{pmatrix} \tag{2.8.2}$$

Proof of Theorem 2.8.2 for scalar value functions on $\mathbb{R}^2$.
Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto f(x, y)$ be a scalar value function on $\mathbb{R}^2$, and let $S \subseteq \mathbb{R}^3$ be the surface described by f. We must prove that a normal vector to the surface S at some point $\mathbf{p} \in \mathbb{R}^2$ is

$$\mathbf{n} = \begin{pmatrix} f_x \\ f_y \\ -1 \end{pmatrix}$$

1. Recall the definition of a normal vector. The vector $\mathbf{n}$ is a normal vector to S at $\mathbf{p}$ $\iff$ $\mathbf{n}$ is orthogonal to all tangent vectors to S at $\mathbf{p}$.
2. By Corrolary 2.7.5, we know that the vectors
$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ f_x \end{pmatrix} \text{ and } \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ f_y \end{pmatrix} \tag{2.8.3}$$
are tangent to the surface S at the point $\mathbf{p}$.
3. Consider $\mathbf{v}_1 \times \mathbf{v}_2$.
$$\begin{aligned} \mathbf{v}_1 \times \mathbf{v}_2 &= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} \\ &= (0 - f_x)\hat{\mathbf{i}} - (f_y - 0)\hat{\mathbf{j}} + (1 - 0)\hat{\mathbf{k}} \\ &= \begin{pmatrix} f_x \\ f_y \\ -1 \end{pmatrix} \\ \implies \mathbf{n} &= \mathbf{v}_1 \times \mathbf{v}_2 \end{aligned} \tag{2.8.4}$$
4. Recall that, by the the definition of the cross product, $\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2$ must be orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.
5. As $\mathbf{v}_1$ and $\mathbf{v}_2$ are coplanar (see 0.3.17) with all other tangent vectors to S at $\mathbf{p}$, $\mathbf{n}$ must be orthogonal to all tangent vectors to S at $\mathbf{p}$.

$\square$