NOTES FOR MATH 376 (SPRING 2018)

STEVEN V SAM

Contents

Date: April 24, 2018 12:23am.

1. Applications of differential calculus

1.1. **Implicit Differentiation.** Let $U \subseteq \mathbf{R}^n$ be an open subset, $F\colon U \to \mathbf{R}$ differentiable. Given open subset $V \subseteq \mathbf{R}^{n-1}$ and $f\colon V \to \mathbf{R}$, say that x_n is **implicitly defined** by f if

$$F(a_1, \dots, a_{n-1}, f(a_1, \dots, a_{n-1})) = 0 \text{ for all } (a_1, \dots, a_{n-1}) \in V.$$

Theorem 1.1. *Suppose f is differentiable. Then for $k = 1, \dots, n-1$ and $(a_1, \dots, a_{n-1}) \in V$, we have*

$$\frac{\partial f}{\partial x_k}(a_1, \dots, a_{n-1}) = -\frac{\partial F/\partial x_k(a_1, \dots, a_{n-1}, f(a_1, \dots, a_{n-1}))}{\partial F/\partial x_n(a_1, \dots, a_{n-1}, f(a_1, \dots, a_{n-1}))}$$

whenever $\partial F/\partial x_n(a_1, \dots, a_{n-1}) \neq 0$.

Proof. Define $g\colon V \to \mathbf{R}$ by

$$g(x_1, \dots, x_{n-1}) = F(x_1, \dots, x_{n-1}, f(x_1, \dots, x_{n-1}))$$

and $h\colon V \to \mathbf{R}^n$ by $(x_1, \dots, x_{n-1}) \mapsto (x_1, \dots, x_{n-1}, f(x_1, \dots, x_{n-1}))$. Then $g = F \circ h$, so by the chain rule we have

$$\begin{bmatrix} \frac{\partial g}{\partial x_1} & \cdots & \frac{\partial g}{\partial x_{n-1}} \end{bmatrix} = \begin{bmatrix} \frac{\partial F}{\partial x_1} & \cdots & \frac{\partial F}{\partial x_n} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 1 \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_{n-1}} \end{bmatrix}$$

Since $g = 0$ on V, the left side is 0, so we get the result by multiplying out the right side. □

Remark 1.2. The *implicit function theorem* gives conditions under which a function f exists. □

Example 1.3. $F(x, y, z) = x^3 + y^3 + z^3 + 6xyz - 1$. Assuming z is implicitly defined, we have

$$\frac{\partial z}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z} = -\frac{3x^2 + 6yz}{3z^2 + 6xy}$$

$$\frac{\partial z}{\partial y} = -\frac{\partial F/\partial y}{\partial F/\partial z} = -\frac{3y^2 + 6xz}{3z^2 + 6xy}.$$

□

A natural generalization is to assume that several of the variables are defined implicitly, i.e., we're given a function $F\colon U \to \mathbf{R}^m$ with $U \subseteq \mathbf{R}^n$ and a function $f\colon V \to \mathbf{R}^m$ with $V \subseteq \mathbf{R}^{n-m}$ such that $F(a_1, \dots, a_{n-m}, f(a_1, \dots, a_{n-m})) = 0$ for all $(a_1, \dots, a_{n-m}) \in V$.

For simplicity, assume $n = 3$ and $m = 2$. Use x, y, z as coordinates and suppose $y = Y(x)$ and $z = Z(x)$ are defined implicitly. Write the function $F\colon U \to \mathbf{R}^2$ as (F_1, F_2). If we use the chain rule as before, we get two equations on V:

$$0 = \frac{\partial F_1}{\partial x} + \frac{\partial F_1}{\partial y} Y'(x) + \frac{\partial F_1}{\partial z} Z'(x)$$
$$0 = \frac{\partial F_2}{\partial x} + \frac{\partial F_2}{\partial y} Y'(x) + \frac{\partial F_2}{\partial z} Z'(x),$$

which we can rewrite in matrix form:

$$\begin{bmatrix} \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{bmatrix} \begin{bmatrix} Y'(x) \\ Z'(x) \end{bmatrix} = \begin{bmatrix} -\frac{\partial F_1}{\partial x} \\ -\frac{\partial F_2}{\partial x} \end{bmatrix}$$

Whenever the determinant of the 2×2 matrix on the left side is nonzero, we can solve for $Y'(x)$ and $Z'(x)$. For example, using Cramer's rule, we get

$$Y'(x) = \frac{\det \begin{bmatrix} -\frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial z} \\ -\frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial z} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{bmatrix}} \qquad Z'(x) = \frac{\det \begin{bmatrix} \frac{\partial F_1}{\partial y} & -\frac{\partial F_1}{\partial x} \\ \frac{\partial F_2}{\partial y} & -\frac{\partial F_2}{\partial x} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{bmatrix}}.$$

1.2. **Second-order Taylor formula for scalar fields.** Let $U \subseteq \mathbf{R}^n$ be an open subset and let $f\colon U \to \mathbf{R}$ be a twice-differentiable function with continuous second partial derivatives. Define the **Hessian** of f to be the $n \times n$ matrix given by

$$H(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{i,j=1}^{n}.$$

Our assumption on f implies that H is a symmetric matrix, i.e., partial derivatives commute.

Theorem 1.4. *Given $a \in U$ and $y \in \mathbf{R}^n$ such that $a + uy \in U$ for all $u \in [-1, 1]$, we have*

(1) *There exists $0 < c < 1$ such that*

$$f(a + y) - f(a) = \nabla f(a) \cdot y + \frac{1}{2} y H(a + cy) y^T. \tag{1.4a}$$

(2) *There exists a function E_2 such that $\lim_{y \to 0} E_2(a, y) = 0$ and*

$$f(a + y) - f(a) = \nabla f(a) \cdot y + \frac{1}{2} y H(a) y^T + \|y\|^2 E_2(a, y). \tag{1.4b}$$

Proof. (1) Define $g\colon [-1, 1] \to \mathbf{R}$ by $g(u) = f(a + uy)$. Then $f(a + y) - f(a) = g(1) - g(0)$. By the Lagrange remainder theorem, there exists $0 < c < 1$ such that

$$g(1) - g(0) = g'(0) + \frac{1}{2} g''(c).$$

If we define $r(u) = a + uy$, then $g = f \circ r$, so we can use the chain rule:

$$\begin{aligned} g'(u) &= \nabla f(r(u)) \cdot r'(u) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(r(u))y_j \\ g'(0) &= \nabla f(a) \cdot y \\ g''(u) &= \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(r(u))y_j = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} f(r(u))y_i y_j = yH(a+uy)y^T. \end{aligned}$$

(2) Let c be as above. Define

$$E_2(a,y) = \begin{cases} \frac{1}{2\|y\|^2} y[H(a+cy) - H(a)]y^T & \text{if } y \neq 0 \\ 0 & \text{if } y = 0 \end{cases}.$$

Using (1.4a), this satisfies (1.4b), so it suffices to show that $\lim_{y \to 0} E_2(a,y) = 0$. We have

$$\begin{aligned} \|y\|^2 |E_2(a,y)| &= \frac{1}{2} \left| \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+cy) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a) y_i y_j \right| \\ &\le \frac{1}{2} \sum_{i,j=1}^{n} \left| \frac{\partial^2 f}{\partial x_i \partial x_j}(a+cy) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right| |y_i y_j| \\ &\le \frac{1}{2} \sum_{i,j=1}^{n} \left| \frac{\partial^2 f}{\partial x_i \partial x_j}(a+cy) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right| \|y\|^2 \end{aligned}$$

since $|y_i y_j| \le \|y\|^2$. In particular,

$$|E_2(a,y)| \le \frac{1}{2} \sum_{i,j=1}^{n} \left| \frac{\partial^2 f}{\partial x_i \partial x_j}(a+cy) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right|,$$

and the result follows since the limit of the terms in the right hand side goes to 0 as $y \to 0$ by continuity. □

1.3. **Maxima, minima, saddle points.** Let $U \subseteq \mathbf{R}^n$ be a subset and let $f \colon U \to \mathbf{R}$ be differentiable. Given $a \in U$, we say that a is:

- an **absolute maximum** if $f(a) \ge f(x)$ for all $x \in U$,
- an **relative maximum** if $f(a) \ge f(x)$ for all x in some ball around a,
- an **absolute minimum** if $f(a) \le f(x)$ for all $x \in U$,
- an **relative minimum** if $f(a) \le f(x)$ for all x in some ball around a.

In all of these cases, call a an **extremum**. All partial derivatives of f are 0 are a: by restricting to a line in each direction, this reduces to the 1-variable case and then follows from single variable calculus. In general, a **critical point** of f is a point where all partial derivatives vanish.

Example 1.5. $f(x,y) = x^4 + y^4 - 4xy$. Its critical points are $(0,0)$, $(1,1)$, $(-1,-1)$. □

Functions in 1 variable can have inflection points (i.e., points which are not relative maxima nor minima), and in the general case, we call such points saddle points: more formally, a is a **saddle point** if it is a critical point, but every n-ball around a contains points x, x' such that $f(a) < f(x)$ and $f(a) > f(x')$.

Example 1.6.
- $f(x, y) = y^2 - x^2$. Then $(0, 0)$ is a saddle point.
- $f(x, y) = x^2 + y^2$. Then $(0, 0)$ is an absolute minimum.
- $f(x, y) = -x^2 - y^2$. Then $(0, 0)$ is an absolute maximum. □

Recall from Math 375 that if A is a real $n \times n$ symmetric matrix, then A is diagonalizable and all of its eigenvalues are real.

Theorem 1.7. *Let A be a real $n \times n$ symmetric matrix. Then:*
1. *$yAy^T > 0$ for all $y \neq 0$ if and only if all eigenvalues of A are positive.*
2. *$yAy^T < 0$ for all $y \neq 0$ if and only if all eigenvalues of A are negative.*

In case (1), A is **positive definite**, and in case (2), A is **negative definite**.

Proof. Let C be an orthogonal matrix such that CAC^T is a diagonal matrix with entries $\lambda_1, \dots, \lambda_n$. Since C is invertible, we conclude that the theorem is true for A if and only if it is true for CAC^T. But

$$yCAC^Ty^T = \sum_{i=1}^n \lambda_i y_i^2,$$

so the result is clear: for example, if all $\lambda_i > 0$, then the right side is always positive if $y \neq 0$, and conversely, if the right hand side is positive for all $y \neq 0$, we conclude that $\lambda_i > 0$ by taking y to be the ith standard basis vector. □

Theorem 1.8. *Let $U \subseteq \mathbf{R}^n$ be an open subset and let $f : U \to \mathbf{R}$ have continuous second partial derivatives. Let $a \in U$ be a critical point of f. Then*
1. *If all eigenvalues of $H(a)$ are positive, then a is a relative minimum of f.*
2. *If all eigenvalues of $H(a)$ are negative, then a is a relative maximum of f.*
3. *If $H(a)$ has positive and negative eigenvalues, then a is a saddle point of f.*

Proof. We just prove (1). Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $H(a)$, which we assume to be positive. Let $h = \min(\lambda_1, \dots, \lambda_n)$ and set $Q(y) = yH(a)y^T$. Then $H(a) - \frac{h}{2}\mathrm{Id}$ has positive eigenvalues and hence is positive definite, so $y(H(a) - \frac{h}{2}\mathrm{Id})y^T > 0$ for all $y \neq 0$, or equivalently, $Q(y) > \frac{h}{2}\|y\|^2$ for all $y \neq 0$.

As in (1.4b) , write (since a is critical, $\nabla f(a) = 0$) out the Taylor formula

$$f(a + y) - f(a) = \frac{1}{2}Q(y) + \|y\|^2 E_2(a, y).$$

where $\lim_{y\to 0} E_2(a, y) = 0$. In particular, there exists $r > 0$ such that $\|y\| < r$ implies that $|E_2(a, y)| < h/4$. So, for such y, we have

$$\|y\|^2 |E_2(a, y)| < \frac{h}{4}\|y\|^2 < \frac{1}{2}Q(y).$$

By the Taylor formula above, we have

$$f(a + y) - f(a) \geq \frac{1}{2}Q(y) - \|y\|^2|E_2(a, y)| > 0$$

whenever $\|y\| < r$, so a is a relative minimum. □

Corollary 1.9. *Let $U \subseteq \mathbf{R}^2$ be an open subset and let $f : U \to \mathbf{R}$ have continuous second partial derivatives and let $a \in U$ be a critical point. Set*

$$A = \frac{\partial^2 f}{\partial x_1 \partial x_1}(a) \qquad B = \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) \qquad C = \frac{\partial^2 f}{\partial x_2 \partial x_2}(a) \qquad \Delta = AC - B^2.$$

Then

(1) $\Delta < 0$ *implies* a *is a saddle point.*
(2) $\Delta > 0$ *and* $A > 0$ *implies* a *is a relative minimum.*
(3) $\Delta > 0$ *and* $A < 0$ *implies* a *is a relative maximum.*

Proof. The Hessian of f at a is the matrix $\begin{bmatrix} A & B \\ B & C \end{bmatrix}$. Let λ, μ be its eigenvalues. Recall that the sum of the eigenvalues of a matrix is its trace and the product of the eigenvalues is its determinant, so we have

$$\lambda + \mu = A + C, \qquad \lambda\mu = \Delta.$$

If $\Delta < 0$, then λ and μ have opposite signs, so a is a saddle point.

Now suppose $\Delta > 0$. Then λ and μ have the same sign, so are both positive or are both negative. Also, $AC > B^2$, so A and C must have the same sign. In particular, if $A > 0$, then $\lambda + \mu > 0$, so they are both positive. Otherwise, if $A < 0$, then $\lambda + \mu < 0$, so they are both negative. □

1.4. **Lagrange multipliers.** The setup for this section: we're given $f, g_1, \dots, g_m \colon \mathbf{R}^n \to \mathbf{R}$, and we want to maximize $f(x_1, \dots, x_n)$ where $(x_1, \dots, x_n)$ satisfies the equations $g_1(x_1, \dots, x_n) = \cdots = g_m(x_1, \dots, x_n) = 0$.

There is a general approach which works in various examples; we'll just explore a few different special cases.

Consider the case $n = 2$ and $m = 1$, so we're trying to maximize the value of $f(x, y)$ along the curve of points (x, y) satisfying $g(x, y) = 0$.

Theorem 1.10. *Suppose that the set of points satisfying $g(x, y) = 0$ can be parametrized by a curve $\alpha \colon \mathbf{R} \to \mathbf{R}^2$ whose derivative is nowhere 0. If (a, b) maximizes f subject to $g(a, b) = 0$, then ∇f and ∇g are parallel at (a, b).*

If $\nabla g(a, b) \neq 0$ whenever $g(a, b) = 0$, then this can be reformulated as saying that there exists $\lambda \in \mathbf{R}$ such that $\nabla f(a, b) = \lambda \nabla g(a, b)$.

We recall one thing from linear algebra: if v is a nonzero vector in $\mathbf{R}^n$, then the set of vectors orthogonal to v is a subspace of dimension $n - 1$ in $\mathbf{R}^n$.

Proof. Define $\varphi \colon \mathbf{R} \to \mathbf{R}$ by $\varphi(t) = f(\alpha(t))$. Our goal is to maximize φ. We know that if this occurs at a value t_0, then $\varphi'(t_0) = 0$. Let α_1, α_2 be the components of α. Use the chain rule:

$$\varphi'(t_0) = \frac{\partial f}{\partial x}(\alpha(t_0))\alpha_1'(t_0) + \frac{\partial f}{\partial y}(\alpha(t_0))\alpha_2'(t_0) = \nabla f(\alpha(t_0)) \cdot \alpha'(t_0).$$

Recall that the gradient of g at each point is orthogonal to the tangent vectors of its level curves[1], so $\nabla f(\alpha(t_0))$ and $\nabla g(\alpha(t_0))$ are both orthogonal to $\alpha'(t_0)$. Hence they are scalar multiples of each other. □

[1] If we instead define $\psi(t) = g(\alpha(t))$, then ψ is identically 0, so the chain rule gives $0 = \nabla g(\alpha(t)) \cdot \alpha'(t)$ for all t.

Example 1.11. Maximize $f(x,y) = x^2 + 2y^2$ subject to the condition $0 = g(x,y) = x^2 + y^2 - 1$. First we set $\nabla f = \lambda \nabla g$:

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x} \Longrightarrow 2x = 2\lambda x$$
$$\frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y} \Longrightarrow 4y = 2\lambda y.$$

The first equation says $(\lambda - 1)x = 0$, so either $\lambda = 1$ or $x = 0$.

- In the first case ($\lambda = 1$), the second equation becomes $4y = 2y$, so $y = 0$. Since $g(x,y) = 0$, this gives $x^2 - 1 = 0$, so $x = \pm 1$. In these cases, we have $f(1,0) = 1$ and $f(-1,0) = 1$.
- In the second case ($x = 0$), the equation $g(x,y) = 0$ forces $y = \pm 1$. In those cases, $f(0,1) = 2$ and $f(0,-1) = 2$.

In particular, $(\pm 1, 0)$ are not actual maxima, but $(0, \pm 1)$ is.

To verify that 2 is the maximum value independently, we can analyze some more. Since $g(x,y) = 0$, we know that $|y| \leq 1$. Also, we can do the substitution $x^2 = -y^2 + 1$ into f, to get $f(x,y) = (-y^2 + 1) + 2y^2 = y^2 + 1$, and since $|y| \leq 1$, this quantity must be ≤ 2. So actually $(0, \pm 1)$ are true maxima. □

We can think of the method of Lagrange multipliers as a way to guess potential maximizing points. In the general case, we could try to prove an analogue of Theorem 1.10 under some conditions on the solution sets of the functions g_i. The outcome would be that ∇f is a linear combination of the ∇g_i, i.e., we have an expression

$$\nabla f = \lambda_1 \nabla g_1 + \cdots + \lambda_m \nabla g_m$$

for some $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$. We won't generalize Theorem 1.10, but let's try this method in one example with $n = 3$ and $m = 2$:

Example 1.12. Maximize $f(x,y,z) = x + 2y + 3z$ subject to the conditions that $x - y + z = 1$ and $x^2 + y^2 = 1$.

Set $g_1(x,y,z) = x - y + z - 1$ and $g_2(x,y,z) = x^2 + y^2 - 1$. Then the method of Lagrange multipliers tells to find solutions to

$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2.$$

This gives 3 equations by considering each partial derivative:

$$1 = \lambda_1 + 2\lambda_2 x$$
$$2 = -\lambda_1 + 2\lambda_2 y$$
$$3 = \lambda_1$$

The first implies that $\lambda_2 x = -1$ and the second implies that $\lambda_2 y = 5/2$. Take the condition $g_2 = 0$ and multiply by λ_2^2:

$$(x\lambda_2)^2 + (y\lambda_2)^2 - \lambda_2^2 = 0 \Longrightarrow 1 + \frac{25}{4} = \lambda_2^2,$$

so $\lambda_2 = \pm\sqrt{29}/2$. Then we can go back and solve for x and y.

- If $\lambda_2 = \sqrt{29}/2$, then $x = -2/\sqrt{29}$, $y = 5/\sqrt{29}$, and using $g_1 = 0$, we get $z = -x + y + 1 = 2/\sqrt{29} + 5/\sqrt{29} + 1 = 1 + 7/\sqrt{29}$. In that case, $f(x,y,z) = 3 + \sqrt{29}$.

- If $\lambda_2 = -\sqrt{29}/2$, we instead get $x = 2/\sqrt{29}$, $y = -5/\sqrt{29}$, and $z = -2/\sqrt{29} - 5/\sqrt{29} + 1 = 1 - 7/\sqrt{29}$. In that case, $f(x, y, z) = 3 - \sqrt{29}$.

So the method of Lagrange multipliers gives $3 + \sqrt{29}$ as the maximum value (we haven't stated a precise theorem when it's valid though). $\square$

2. Line integrals

2.1. **Definitions.** Given real numbers $a < b$, a function $\alpha\colon [a, b] \to \mathbf{R}^n$ is a

- **smooth path** if α' exists and is continuous on (a, b).
- **piecewise smooth path** if there exist $a = a_0 < a_1 < \cdots < a_r = b$ such that the restriction of α to each $[a_i, a_{i+1}]$ is a smooth path.

If α is piecewise smooth and $f\colon \mathbf{R}^n \to \mathbf{R}^n$ is a vector field, we define the **line integral of** f **along** α to be

$$\int f \cdot \mathrm{d}\alpha := \int_a^b f(\alpha(t)) \cdot \alpha'(t)\mathrm{d}t$$

whenever the right side exists. If C is the image of α, the integral is also denoted $\int_C f \mathrm{d}\alpha$.

Example 2.1. Let $a = 0, b = 1$, $n = 2$, and $\alpha(t) = (t, t^2)$, and $f(x, y) = (x^2 + y^2, x^2 - y^2)$. Then $\alpha'(t) = (1, 2t)$, so

$$\int f \cdot \mathrm{d}\alpha = \int_0^1 (t^2 + t^4, t^2 - t^4) \cdot (1, 2t)\mathrm{d}t = \int_0^1 (t^2 + t^4 + 2t^3 - 2t^5)\mathrm{d}t. \qquad \square$$

2.2. **Basic properties.** Given a path C which is the union of two paths C_1 and C_2, we write $C = C_1 + C_2$. Then for $a, b \in \mathbf{R}$ and vector fields $f, g\colon \mathbf{R}^n \to \mathbf{R}^n$, we have the following two additivity properties:

- $\displaystyle\int_C (af + bg) \cdot \mathrm{d}\alpha = a \int_C f \cdot \mathrm{d}\alpha + b \int_C g \cdot \mathrm{d}\alpha.$
- $\displaystyle\int_{C_1 + C_2} f \cdot \mathrm{d}\alpha = \int_{C_1} f \cdot \mathrm{d}\alpha + \int_{C_2} f \cdot \mathrm{d}\alpha.$

Let $\alpha\colon [a, b] \to \mathbf{R}^n$ be a piecewise smooth path and let $u\colon [c, d] \to [a, b]$ be a surjective continuously differentiable function such that u' is nowhere 0. Define $\beta\colon [c, d] \to \mathbf{R}^n$ by $\beta(t) = \alpha(u(t))$. We say that α and β are **equivalent**. Furthermore, if $u' > 0$ on $[c, d]$, then u is **orientation-preserving** and α and β go in the **same direction**. If $u' < 0$ on $[c, d]$, then u is **orientation-reversing** and α and β go in **opposite directions**.

Theorem 2.2. *Let α and β be equivalent piecewise smooth paths.*

- *If α and β go in the same direction, then* $\displaystyle\int f \cdot \mathrm{d}\alpha = \int f \cdot \mathrm{d}\beta.$
- *If α and β go in opposite directions, then* $\displaystyle\int f \cdot \mathrm{d}\alpha = -\int f \cdot \mathrm{d}\beta.$

Proof. By additivity, we can reduce to the case that α and β are equivalent smooth paths. Use the notation above, so that $\beta = \alpha \circ u$. Set $v = u(t)$. Then we have

$$\begin{aligned}\int f \cdot \mathrm{d}\beta &= \int_c^d f(\beta(t)) \cdot \beta'(t)\mathrm{d}t = \int_c^d f(\alpha(u(t))) \cdot \alpha'(u(t))u'(t)\mathrm{d}t \\ &= \int_{u(c)}^{u(d)} f(\alpha(v)) \cdot \alpha'(v)\mathrm{d}v.\end{aligned}$$

If $u' > 0$, then $u(c) = a$ and $u(d) = b$, so the last integral is $\int f \cdot \mathrm{d}\alpha$. Otherwise, they are swapped, and the last integral is $-\int f \cdot \mathrm{d}\alpha$. □

2.3. **Connected sets and path independence.** Let $S \subseteq \mathbf{R}^n$ be an open subset. We say that S is **disconnected** if we can write $S = S_1 \cup S_2$ where S_1 and S_2 are disjoint and nonempty open subsets. If S is not disconnected, then it is **connected**.

We will use the following fact: If S is connected and $a, b \in S$, then there is a continuous function $\alpha\colon [0,1] \to S$ such that $\alpha(0) = a$ and $\alpha(1) = b$, i.e., any two points can be joined together with a path that stays inside of S.

Let $f\colon S \to \mathbf{R}^n$ be a continuous vector field. Given $a, b \in S$, we say that the integral of f is **independent of the path from a to b** if the value of the integral $\int f \cdot \mathrm{d}\alpha$ is always the same for any path α that starts at a and ends at b. The integral of f is **independent of the path in** S if it is independent of the path from a to b for all choices of points a, b.

Example 2.3. (1) Consider $S = \mathbf{R}^2$ and $f(x,y) = (\sqrt{y}, x^3 + y)$. Take $a = (0,0)$ and $b = (1,1)$. For $\alpha\colon [0,1] \to \mathbf{R}^2$ given by $\alpha(t) = (t,t)$, we have $\int f \cdot \mathrm{d}\alpha = 17/12$, and for $\beta\colon [0,1] \to \mathbf{R}^2$ given by $\beta(t) = (t^2, t^3)$, we have $\int f \cdot \mathrm{d}\beta = 59/42$, so the integral of f is not independent of the path.

(2) Consider instead $f(x,y) = (x,y)$. Then for any curve $\alpha\colon [a,b] \to \mathbf{R}^2$, we have

$$\int f \cdot \mathrm{d}\alpha = \int_a^b f(\alpha(t)) \cdot \alpha'(t)\mathrm{d}t = \int_a^b (\alpha_1(t)\alpha_1'(t) + \alpha_2(t)\alpha_2'(t))\mathrm{d}t = \frac{1}{2}(\alpha_1^2(t) + \alpha_2^2(t))|_a^b,$$

where the last equality comes from a u-substitution. The last quantity only depends on the endpoints $\alpha(a)$ and $\alpha(b)$, but not on the actual path α, so the integral of f is independent of path. □

2.4. **Fundamental theorems of calculus for line integrals.** We will use the following version of the fundamental theorem of calculus from single-variable calculus:

Theorem 2.4 (Second fundamental theorem of calculus). *Let $\varphi\colon [a,b] \to \mathbf{R}$ be a continuous function such that φ' exists and is continuous on (a,b). Assuming that $\int_a^b \varphi'(t)\mathrm{d}t$ is defined, we have*

$$\int_a^b \varphi'(t)\mathrm{d}t = \varphi(b) - \varphi(a).$$

Here is the analogue for line integrals:

Theorem 2.5 (Second fundamental theorem of calculus for line integrals). *Let $U \subseteq \mathbf{R}^n$ be an open connected subset and let $\varphi\colon U \to \mathbf{R}$ be a differentiable function whose gradient $\nabla\varphi$ is continuous. Given a piecewise smooth path $\alpha\colon [a,b] \to U$, we have*

$$\int \nabla\varphi \cdot \mathrm{d}\alpha = \varphi(\alpha(b)) - \varphi(\alpha(a)).$$

Proof. First assume that α is smooth, not just piecewise smooth. Define $g\colon [a,b] \to \mathbf{R}$ by $g(t) = \varphi(\alpha(t))$. By the chain rule, we have

$$g'(t) = \nabla\varphi(\alpha(t)) \cdot \alpha'(t).$$

Since $\alpha'(t)$ is continuous on (a,b) and $\nabla\varphi$ is continuous on U, we conclude that $g'(t)$ is also continuous on (a,b). This, together with the definition of the line integral, gives us

$$\int \nabla\varphi \cdot \mathrm{d}\alpha = \int_a^b \nabla\varphi(\alpha(t)) \cdot \alpha'(t)\mathrm{d}t = \int_a^b g'(t)\mathrm{d}t = g(b) - g(a) = \varphi(\alpha(b)) - \varphi(\alpha(a)).$$

where the second to last equality is Theorem 2.4.

Now consider the general case that α is only piecewise smooth. Then we can find $a = a_0 < a_1 < \cdots < a_r = b$ so that α is smooth on each $[a_i, a_{i+1}]$. Let α^i be the restriction of α to $[a_i, a_{i+1}]$. By additivity, we have

$$\int \nabla\varphi \cdot \mathrm{d}\alpha = \sum_{i=0}^{r-1} \int \nabla\varphi \cdot \mathrm{d}\alpha^i = \sum_{i=0}^{r-1} (\varphi(\alpha^i(a_{i+1})) - \varphi(\alpha^i(a_i))) = \varphi(\alpha(b)) - \varphi(\alpha(a)).$$

For the last equality, we use that α^i agrees with α on the interval $[a_i, a_{i+1}]$; then we simplify the telescoping sum. □

We get the following consequence: the integral of the gradient of a scalar field is independent of the path. In particular, if $\alpha(a) = \alpha(b)$, i.e., α traces out a closed loop, then the integral is 0.

Now we discuss the analogue of the first fundamental theorem of calculus. Let $U \subseteq \mathbf{R}^n$ be an open connected set and let $f \colon U \to \mathbf{R}^n$ be a vector field such that the integral of f is independent of the path in U. Fix a point $a \in U$. Then for any other point $x \in U$, the value of $\int f \cdot \mathrm{d}\alpha$ is independent of the choice of a path α starting at a and ending at x, so we can denote it by $\int_a^x f \cdot \mathrm{d}\alpha$ to be more suggestive with the notation.

Define a function $\varphi \colon U \to \mathbf{R}$ by

$$\varphi(x) = \int_a^x f \cdot \mathrm{d}\alpha.$$

Theorem 2.6 (First fundamental theorem of calculus for line integrals)**.** *With notation as above, the gradient of φ exists and $\nabla\varphi(x) = f(x)$ for all $x \in U$.*

Proof. Fix $x \in U$. Let e_k be a coordinate unit vector and let $(f_1, \dots, f_n)$ be the component functions of f. Since U is open, there is a ball of radius $r > 0$ around x contained in U, so $x + he_k \in U$ for $h \in [-r, r]$. We want to show that

$$\lim_{h \to 0} \frac{\varphi(x + he_k) - \varphi(x)}{h}$$

exists, and is equal to f_k. The numerator can be rewritten as

$$\varphi(x + he_k) - \varphi(x) = \int_a^{x+he_k} f \cdot \mathrm{d}\alpha - \int_a^x f \cdot \mathrm{d}\alpha = \int_x^{x+he_k} f \cdot \mathrm{d}\alpha.$$

where we used additivity of line integrals and Theorem 2.2. By our assumption, the last integral can be computed using any path α from x to $x + he_k$, so we'll use $\alpha \colon [0, 1] \to U$ given by $\alpha(t) = x + the_k$. Then we have

$$\frac{\varphi(x + he_k) - \varphi(x)}{h} = \frac{1}{h} \int_x^{x+he_k} f \cdot \mathrm{d}\alpha = \frac{1}{h} \int_0^1 f(x + the_k) \cdot he_k \mathrm{d}t = \int_0^1 f_k(x + the_k) \mathrm{d}t.$$

Now do a change of variables: $u = ht$, $\mathrm{d}u = h\mathrm{d}t$ to get

$$\frac{1}{h} \int_0^h f_k(x + ue_k) \mathrm{d}u.$$

Next, define $g\colon [-r, r] \to \mathbf{R}$ by

$$g(t) = \int_0^t f_k(x + ue_k)\mathrm{d}u.$$

Since f_k is continuous, we can apply the usual fundamental theorem of calculus to conclude that $g'(t) = f_k(x + te_k)$. Combining everything:

$$\lim_{h\to 0} \frac{\varphi(x + he_k) - \varphi(x)}{h} = \lim_{h\to 0} \frac{1}{h}\int_0^h f_k(x + ue_k)\mathrm{d}u = \lim_{h\to 0} \frac{g(h) - g(0)}{h} = g'(0) = f_k(x). \quad \square$$

2.5. **When is a vector field a gradient?** Given a vector field f, we want to know when it is the gradient of a differentiable function φ. If φ exists, we call it a **potential function**. The fundamental theorems of calculus for line integrals give us equivalent conditions for this to be true. Call a path α **closed** if its starting and end points are the same.

Theorem 2.7. *Let $U \subseteq \mathbf{R}^n$ be an open connected subset and let $f\colon U \to \mathbf{R}^n$ be a continuous vector field. Then the following are equivalent:*

(a) *f is the gradient of a potential function.*
(b) *The line integral of f is independent of the path in U.*
(c) *The line integral of f is 0 around every piecewise smooth closed path in U.*

Proof. We show (a) $\implies$ (c) $\implies$ (b) $\implies$ (a).

(a $\implies$ c) First suppose (a) holds, so f is the gradient of a function φ. If α is a piecewise smooth closed path, then the second fundamental theorem implies that the integral of f around α is $\varphi(x) - \varphi(x) = 0$, so (c) holds.

(c $\implies$ b) Now suppose that (c) holds. Let α and β be any two piecewise smooth closed path with the same starting and end points. Let γ be the closed loop which first does α, and then goes backwards along β. By (c), the integral of f along γ is 0. By additivity, we have

$$\int f \cdot \mathrm{d}\gamma = \int f \cdot \mathrm{d}\alpha - \int f \cdot \mathrm{d}\beta,$$

so we conclude that the integral along α and β are the same, which shows (b) holds.

(b $\implies$ a) If (b) holds, we can use the first fundamental theorem to construct the potential function. $\square$

Example 2.8. Going back to Example 2.3, let $f(x, y) = (\sqrt{y}, x^3 + y)$. We found two different paths from $(0, 0)$ to $(1, 1)$ such that the integral of f along these paths give different values. In particular, we conclude that f is *not* the gradient of a function. $\square$

We can also give a necessary condition for f to be the gradient of a function which does not use integrals:

Theorem 2.9. *Let $f = (f_1, \dots, f_n)$ be a continuously differentiable vector field on an open subset $U \subseteq \mathbf{R}^n$. If f is the gradient of a potential function φ, then*

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i} \text{ for all } i, j.$$

Proof. This is just the statement that if φ has continuous second derivatives, then its second partial derivatives can be computed in either order. $\square$

Example 2.10. Continue with the previous example $f(x,y) = (\sqrt{y}, x^3 + y)$. Then

$$\frac{\partial f_1}{\partial y} = \frac{1}{2\sqrt{y}}, \qquad \frac{\partial f_2}{\partial x} = 3x^2,$$

so this gives another reason why f is not the gradient of a potential function. $\square$

Example 2.11. It is important to note that the criteria using derivatives is a necessary condition for f to be a gradient, but not sufficient. For example, take $U = \mathbf{R}^2 \setminus \{(0,0)\}$ and

$$f(x,y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right).$$

Then

$$\frac{\partial f_1}{\partial y} = \frac{-x^2+y^2}{(x^2+y^2)^2} = \frac{\partial f_2}{\partial x}.$$

However, we claim that f is *not* the gradient of a potential function. Let $\alpha \colon [0, 2\pi] \to U$ be the closed path $\alpha(t) = (\cos t, \sin t)$. Then

$$\int f \cdot \mathrm{d}\alpha = \int_0^{2\pi} (-\sin t, \cos t) \cdot (-\sin t, \cos t) \mathrm{d}t = \int_0^{2\pi} \mathrm{d}t = 2\pi.$$

Since it is nonzero, Theorem 2.7 tells us f is not a gradient. $\square$

Remark 2.12. In the example above, U has a "hole": we've removed the point $(0,0)$. Roughly speaking, if U does not have holes, then the necessary condition with derivatives is also a sufficient condition. In general, the failure of this to be true can be detected by the first de Rham cohomology group of U, $\mathrm{H}^1_{\mathrm{dR}}(U)$. We might touch on this topic later in the course if there is time. $\square$

2.6. **Convex regions.** A subset $U \subseteq \mathbf{R}^n$ is **convex**, if given any two points $a, b \in U$, the line segment from a to b is contained in U. Given a vector field $f \colon U \to \mathbf{R}^n$, Apostol shows that $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ for all i, j implies that f is the gradient of a potential function. In fact, the proof implies a more general statement, so we state that version here:

Theorem 2.13. *Let $U \subseteq \mathbf{R}^n$ be an open connected subset and assume that there is a point $a \in U$ such that for all other $x \in U$, the line segment between a and x is contained in U. Given a continuously differentiable vector field $f \colon U \to \mathbf{R}^n$, we have that f is the gradient of a potential function if and only if $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ for all i, j.*

The proof of this theorem requires a technical statement (whose proof we omit, but can be found in §10.21).

Call a product of intervals $[a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbf{R}^n$ an **interval** (slightly confusing, though we're following Apostol's terminology here). It has nonempty interior if $b_i > a_i$ for all i.

Theorem 2.14. *Let S be an interval in $\mathbf{R}^n$ with nonempty interior and let $J = [a, b]$. Write points of $S \times J$ as (x, t) where $x \in S$ and $t \in J$. Let $\psi \colon S \times J \to \mathbf{R}$ be a scalar field such that $\frac{\partial \psi}{\partial x_k}$ is continuous on $S \times J$ for some $k = 1, \ldots, n$. Define $\varphi \colon S \to \mathbf{R}$ by*

$$\varphi(x) = \int_a^b \psi(x, t)\, \mathrm{d}t.$$

For each interior point $x \in S$, we have

$$\frac{\partial \varphi}{\partial x_k} = \int_a^b \frac{\partial \psi}{\partial x_k}(x,t)\,\mathrm{d}t.$$

In other words, we can differentiate under the integral sign.

Proof of Theorem 2.13. We have already seen that if f is a gradient, then $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ for all i, j.

Now suppose that $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ for all i, j. We will use these equalities to construct a potential function for f. Without loss of generality, we may assume that the point a is the origin by translating everything by a. Define $\psi \colon U \times [0,1] \to \mathbf{R}$ by $\psi(x,t) = f(tx) \cdot x$. We define a function $\varphi \colon U \to \mathbf{R}$ by

$$\varphi(x) = \int_0^1 \psi(x,t)\,\mathrm{d}t.$$

(Note that this is just the line integral of a straight line path from the origin to x.) Since U is open, we can find a ball of positive radius centered at x. Inside of that ball, we can find an interval S with nonempty interior. Since f is continuous differentiable, we conclude that ψ is also continuously differentiable on $S \times [0,1]$. Then Theorem 2.14 implies that

$$\frac{\partial \varphi}{\partial x_k}(x) = \int_0^1 \frac{\partial \psi}{\partial x_k}(x,t)\,\mathrm{d}t.$$

Expanding $\psi(x,t) = f_1(tx)x_1 + \cdots + f_n(tx)x_n$, we get

$$\begin{aligned}
\frac{\partial \psi}{\partial x_k}(x,t) &= t\left(\frac{\partial f_1}{\partial x_k}(tx)x_1 + \cdots + \frac{\partial f_n}{\partial x_k}(tx)x_n\right) + f_k(tx) \\
&= t\left(\frac{\partial f_k}{\partial x_1}(tx)x_1 + \cdots + \frac{\partial f_k}{\partial x_n}(tx)x_n\right) + f_k(tx) \\
&= t\nabla f_k(tx) \cdot x + f_k(tx),
\end{aligned}$$

where in the second equality, we used our assumption on the derivatives of f. Now define $g(t) = f_k(tx)$, so that we get

$$\frac{\partial \psi}{\partial x_k}(x,t) = tg'(t) + g(t) = \frac{\mathrm{d}}{\mathrm{d}t}(tg(t)),$$

and in particular,

$$\frac{\partial \varphi}{\partial x_k}(x) = \int_0^1 \frac{\partial \psi}{\partial x_k}(x,t)\,\mathrm{d}t = tg(t)|_0^1 = g(1) = f_k(x).$$

In conclusion, $\nabla \varphi = f$, so we have constructed the desired potential function. $\square$

This gives an explicit way to construct a potential function for a vector field f once we know that $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ for all i, j. In the above proof, we assumed that $a = 0$, and obtained the following formula for the potential:

$$\varphi(x) = \int_0^1 f(tx) \cdot x\,\mathrm{d}t.$$

In general, we would get

$$\varphi(x) = \int_0^1 f(a + t(x-a)) \cdot (x-a)\,\mathrm{d}t.$$

Example 2.15. Consider $U = \mathbf{R}^2$ and $f(x,y) = (x,y)$. Then $\frac{\partial f_1}{\partial y} = 0 = \frac{\partial f_2}{\partial x}$. We can construct a potential function by

$$\varphi(x) = \int_0^1 f(tx) \cdot x \, \mathrm{d}t = \int_0^1 (tx, ty) \cdot (x,y) \, \mathrm{d}t = \int_0^1 t(x^2+y^2) \mathrm{d}t = \frac{1}{2}(x^2+y^2). \qquad \square$$

3. Multiple integrals

Given a region $Q \subseteq \mathbf{R}^n$, and a function $f \colon Q \to \mathbf{R}$, our goal is to define the integral of f over Q, denoted $\iint_Q f$. There is a big difference between $n = 1$ (which you have seen) and $n > 1$. For simplicity, we will stick with $n = 2$. It will usually be clear how to generalize to higher dimensions, but the notation gets more involved.

3.1. **Step functions and their integrals.** First we treat the case that Q is a rectangle, i.e., Q is a product of intervals $Q = [a,b] \times [c,d] = \{(x,y) \in \mathbf{R}^2 \mid a \le x \le b,\ c \le y \le d\}$. We do allow the possibility that $a = b$ or $c = d$ in our definition of rectangle. Given $a = a_0 < a_1 < \cdots < a_r = b$ and $c = c_0 < c_1 < \cdots < c_s = d$, we can subdivide the intervals $[a,b]$ and $[c,d]$ into r and s many intervals, respectively, which in turn subdivide Q into rs smaller rectangles.

Given two subdivisions of an interval, we get a finer subdivision by taking together all of the values from both of them. Similarly, given two subdivisions of a rectangle, we get a common refinement of both of them.

A function $f \colon Q \to \mathbf{R}$ is a **step function** if there exists a refinement of Q into subrectangles such that f is constant on each of the small rectangles.

Lemma 3.1. *Let f, g be step functions on Q and $c_1, c_2 \in \mathbf{R}$. Then $c_1 f + c_2 g$ is also a step function.*

Proof. It follows from the definition that $c_1 f$ and $c_2 g$ are step functions, so it suffices to check that the sum of two step functions is again a step function. To do this, we take the subdivisions of Q on which $c_1 f$ and $c_2 g$ are constant, and take their common refinement to get one where the sum is also constant. $\square$

Keep the notation above, so f is a step function. For $1 \le i \le r$ and $1 \le j \le s$, let Q_{ij} denote the subrectangle $[a_{i-1}, a_i] \times [c_{j-1}, c_j]$ and let $f(Q_{ij})$ denote the common value f takes on any point of that rectangle. We define the integral of f as

$$\iint_Q f = \sum_{i=1}^r \sum_{j=1}^s f(Q_{ij})(a_i - a_{i-1})(c_j - c_{j-1}).$$

In principle, this definition may depend on the subdivision. In homework you will show that it does not. Here are some basic properties:

Theorem 3.2. *Let f, g be step functions on Q.*

(a) *Linearity: For $c_1, c_2 \in \mathbf{R}$, we have*

$$\iint_Q (c_1 f + c_2 g) = c_1 \iint_Q f + c_2 \iint_Q g.$$

(b) *Additivity: If Q is subdivided into two subrectangles Q_1 and Q_2, then*

$$\iint_Q f = \iint_{Q_1} f + \iint_{Q_2} f.$$

(c) *Comparison: If $f(x) \leq g(x)$ for all $x \in Q$, then*

$$\iint_Q f \leq \iint_Q g.$$

3.2. **Integrals of bounded functions.** As before, let Q be a rectangle. Now let f be a bounded function on Q. This means that there is some real number C so that $|f(x)| \leq C$ for all $x \in Q$. If g is another function on Q, we write $g \leq f$ to mean that $g(x) \leq f(x)$ for all $x \in Q$. Define two sets:

$$S = \left\{ \iint_Q s \mid s \text{ is a step function and } s \leq f \right\}$$

$$T = \left\{ \iint_Q t \mid t \text{ is a step function and } t \geq f \right\}.$$

Since f is bounded, there exist step functions $s \leq f$ and also $t \geq f$, so both S and T are nonempty. Note that by the comparison property, for any $\sigma \in S$ and $\tau \in T$, we have $\sigma \leq \tau$. So S is bounded from above, and hence $\sup S$ is well-defined, and similarly, T is bounded from below, and hence $\inf T$ is well-defined and we have $\sup S \leq \inf T$.

The number $\sup S$ is the **lower integral** of f, and is denoted by $\underline{I}(f)$. Similarly, $\inf T$ is the **upper integral** of f, and is denoted by $\overline{I}(f)$. If $\underline{I}(f) = \overline{I}(f)$, then f is called **integrable**, and $\iint_Q f$ is defined to be this common value.

3.3. **Double integrals as repeated one-dimensional integration.** So far, we don't have a way to evaluate multiple integrals in an easy way. In some cases, we can reduce it to computing one-dimensional integrals.

Lemma 3.3. *Let f be a step function on a rectangle $Q = [a, b] \times [c, d]$. Then*

$$\iint_Q f = \int_c^d \left(\int_a^b f(x, y) \mathrm{d}x \right) \mathrm{d}y.$$

Proof. First assume that f is constant on all of Q. Then $\iint_Q f = f(Q)(b-a)(d-c)$. On the other hand,

$$\int_c^d \left(\int_a^b f(x, y) \mathrm{d}x \right) \mathrm{d}y = \int_c^d f(Q)(b-a) \mathrm{d}y = f(Q)(d-c)(b-a).$$

For the general case, we can repeatedly make use of the additivity property for $\iint_Q f$ with respect to subdividing Q as well as a similar additivity property for the repeated integral on the right side of the desired equation. $\square$

Theorem 3.4. *Let f be a bounded, integrable function on $Q = [a, b] \times [c, d]$. For each $y \in [c, d]$, assume that $A(y) = \int_a^b f(x, y) \mathrm{d}x$ exists. If $\int_c^d A(y) \mathrm{d}y$ also exists, then it is equal to $\iint_Q f$, so we have the equality*

$$\iint_Q f = \int_c^d \left(\int_a^b f(x, y) \mathrm{d}x \right) \mathrm{d}y.$$

Proof. Choose step functions s, t on Q such that $s \leq f \leq t$. Integrating over $[a, b]$ with respect to x, we get

$$\int_a^b s(x, y) \mathrm{d}x \leq A(y) \leq \int_a^b t(x, y) \mathrm{d}x.$$

Now we can integrate over $[c,d]$ with respect to y. Using Lemma 3.3, we get

$$\iint_Q s \le \int_c^d A(y)\mathrm{d}y \le \iint_Q t.$$

Since this is valid for all $s \le f$ and all $t \ge f$, we conclude that

$$\underline{I}(f) \le \int_c^d A(y)\mathrm{d}y \le \overline{I}(f).$$

Since f is integrable, the outside two quantities are equal to each other, the common value being $\iint_Q f$. Hence the middle term is also equal to $\iint_Q f$. □

Remark 3.5. Here we integrated with respect to the first variable and then the second variable. There's no particular reason to prefer that, so we can swap the roles of x and y in the statement of the previous theorem. □

Example 3.6. Let $Q = [0,2] \times [1,2]$ and $f(x,y) = x - 3y^2$. Then

$$\begin{aligned}\int_1^2 \left(\int_0^2 (x-3y^2)\mathrm{d}x\right)\mathrm{d}y &= \int_1^2 \left[\frac{x^2}{2} - 3xy^2\right]_0^2 \mathrm{d}y \\ &= \int_1^2 (2-6y^2)\mathrm{d}y \\ &= \left[2y - 2y^3\right]_1^2 = -12.\end{aligned}$$

□

In the 1-variable case, we can interpret the integral $\int_a^b f(x)\,\mathrm{d}x$ as the area under the graph of $f(x)$ (suitably interpreted if f takes negative values). If we apply this to our repeated integration formula, then we can interpret $\iint_Q f$ as the volume of the region under the graph of $f(x,y)$ (again suitably interpreted if f takes negative values). To be precise, first suppose that $f(x,y) \ge 0$ for all $(x,y) \in Q$. We have

$$\iint_Q f = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}x\right)\mathrm{d}y,$$

and the integrand $A(y) = \int_a^b f(x,y)\,\mathrm{d}x$ can be interpreted as the area of the cross-sections of the region under the graph of f. By integrating it once more with respect to y, we get the volume. In the general case, the integral is computing the volume of the region above the xy-plane and below the graph of f minus the region which is below the xy-plane and above the graph of f.

Example 3.7. Let $Q = [-1,1] \times [-3,3]$ and $f(x,y) = \sqrt{1-x^2}$. We can interpret the integral $\iint_Q f$ as the volume of half of a cylinder with radius 1 and height 6, so it is 3π. □

Example 3.8. Let S be the solid which is below the graph of $f(x,y) = 16 - x^2 - 2y^2$ and bounded by the coordinate planes and the planes $x = 2$ and $y = 2$. In this case, it is not obvious how to compute the volume of S, so we can instead set it up as a double integral (set $Q = [0, 2] \times [0, 2]$):

$$\begin{aligned}\mathrm{vol}(S) &= \int_Q (16 - x^2 - 2y^2) \\ &= \int_0^2 \left(\int_0^2 (16 - x^2 - 2y^2)\, \mathrm{d}x \right) \mathrm{d}y \\ &= \int_0^2 \left(\frac{88}{3} - 4y^2 \right) \mathrm{d}y \\ &= 48.\end{aligned}$$

Here we've implicitly assumed that the double integral can be set up as an iterated integral, but there are no problems verifying the hypotheses of Theorem 3.4 since at all steps we're just dealing with polynomial functions. □

3.4. **Integrability of continuous functions.** Our goal is the following theorem:

Theorem 3.9. *Let f be a continuous function on a rectangle $Q = [a, b] \times [c, d]$. Then f is integrable on Q, and moreover, we have*

$$\iint_Q f = \int_c^d \left(\int_a^b f(x, y)\, \mathrm{d}x \right) \mathrm{d}y = \int_a^b \left(\int_c^d f(x, y)\, \mathrm{d}y \right) \mathrm{d}x.$$

In order to prove this, we need some properties of continuous functions on rectangles which we won't prove in this class (it could be done, but might be more appropriate for Math 521).

Theorem 3.10. *Let f be a continuous function on a rectangle Q.*

(a) *f is bounded, i.e., there exists C so that $|f(x)| \le C$ for all $x \in Q$.*
(b) *For any $\varepsilon > 0$, we can subdivide Q into finitely many rectangles $Q_1, \ldots, Q_n$ with the following property: for each i, let $m_i(f)$ be the minimum value of f on Q_i, and let $M_i(f)$ be the maximum value. Then $M_i(f) - m_i(f) < \varepsilon$ for all i.*

Proof of Theorem 3.9. By Theorem 3.10(a), f is bounded, so we can define its upper and lower integrals $\overline{I}(f)$ and $\underline{I}(f)$, and our goal is to show that they are equal. To do this, it suffices to show that the difference $\overline{I}(f) - \underline{I}(f)$ is smaller than any positive number ε.

Let A be the area of Q. By Theorem 3.10(b), we can subdivide Q into finitely many rectangles $Q_1, \ldots, Q_n$ so that, for each i, we have $M_i(f) - m_i(f) < \frac{\varepsilon}{A}$, where $M_i(f)$ is the maximum value f takes on Q_i, and $m_i(f)$ is the minimum value that it takes.

Define step functions s, t on Q as follows: if x is in the interior of Q_i, define $s(x) = m_i(f)$ and $t(x) = M_i(f)$. On the points x of overlap between the Q_i, we define $s(x) = m$ and $t(x) = M$ where $m = \min(m_1, \ldots, m_n)$ and $M = \max(M_1, \ldots, M_n)$. By construction, we have $s \le f \le t$. Also,

$$\iint_Q s = \sum_{i=1}^n m_i(f) \mathrm{area}(Q_i), \qquad \iint_Q t = \sum_{i=1}^n M_i(f) \mathrm{area}(Q_i).$$

In particular, we have

$$\iint_Q s \le \underline{I}(f) \le \overline{I}(f) \le \iint_Q t,$$

and hence

$$\overline{I}(f) - \underline{I}(f) \le \iint_Q t - \iint_Q s = \sum_{i=1}^n (M_i(f) - m_i(f))\mathrm{area}(Q_i) < \frac{\varepsilon}{A}\sum_{i=1}^n \mathrm{area}(Q_i) = \varepsilon.$$

Since we can prove this for any ε, we conclude that $\overline{I}(f) = \underline{I}(f)$, and hence f is integrable.

Finally, we have to prove that $\iint_Q f$ can be evaluated as an iterated integral, so we use Theorem 3.4. There are two iterated integrals, but the proofs are the same in both cases, so we just explain the first one.

First, the integral $A(y) = \int_a^b f(x,y)\,\mathrm{d}x$ exists for all y since f is continuous, and you have already seen that continuous functions are integrable in the 1-variable case. To show that $\int_c^d A(y)\,\mathrm{d}y$ also exists, it suffices to show that $A(y)$ is a continuous function on $[c,d]$. To show this, fix a point $y_0 \in [c,d]$. Then for any other $y_1 \in [c,d]$, pick $x_0 \in [a,b]$ so that $|f(x_0,y_0) - f(x_0,y_1)|$ is maximized amongst all $x \in [a,b]$. Then we have

$$\begin{aligned} |A(y_0) - A(y_1)| &= \left|\int_a^b (f(x,y_0) - f(x,y_1))\,\mathrm{d}x\right| \\ &\le \int_a^b |f(x,y_0) - f(x,y_1)|\,\mathrm{d}x \\ &\le (b-a)|f(x_0,y_0) - f(x_0,y_1)|. \end{aligned}$$

We claim that as y_1 gets closer to y_0, the last quantity can be made arbitrarily small. Pick $\varepsilon > 0$. By Theorem 3.10, we can subdivide Q into finitely many rectangles so that the difference between the min and max values of f on each is less than $\varepsilon/(b-a)$. If $|y_1 - y_0|$ is small enough, then for any x, the points (x,y_0) and (x,y_1) will be in the same rectangle. So in particular, the last quantity above is $< \varepsilon$, which shows that A is continuous at y_0, and finishes the proof. □

3.5. **Double integrals over more general regions.** We'd like to have a definition of integrals over regions which are not necessarily rectangles. First, we need a definition: if A is a bounded subset of the plane (i.e., we can enclose it in a circle of some finite radius), then A has **content zero** if, for every $\varepsilon > 0$, there is a finite set of rectangles whose union contains A and whose total area is at most ε.

Theorem 3.11. *Let f be a bounded function on a rectangle $Q = [a,b] \times [c,d]$. Let D be the set of points where f is not continuous and assume that D has content 0. Then f is integrable on Q.*

Proof. Let C be a number such that $|f(x)| \le C$ for all $x \in Q$. Pick $\delta > 0$. Since D has content 0, we can cover it with a finite number of rectangles whose total area is at most δ. Let R be the union of these rectangles. We can extend this set of rectangles to a partition of Q into subrectangles. Now pick $\varepsilon > 0$. By Theorem 3.10, we can further refine our partition into a new partition so that the difference between the max and min values of f on all rectangles not in R is at most ε.

Define step functions s,t as follows. If x is in the interior of a rectangle Q' not in R, then $s(x)$ is the minimum value that f takes in Q', and $t(x)$ is the maximum value that f takes in Q'. For all other points, set $s(x) = -C$ and $t(x) = C$. Then $s \le f \le t$ and

$$\overline{I}(f) - \underline{I}(f) \le \iint_Q t - \iint_Q s \le \varepsilon(\mathrm{area}(Q) - \mathrm{area}(R)) + 2C\mathrm{area}(R) \le \varepsilon\mathrm{area}(Q) + 2C\delta.$$

Since this is true for any $\varepsilon > 0$, we conclude that $\overline{I}(f) - \underline{I}(f) \leq 2C\delta$. However, this latter inequality is also valid for any $\delta > 0$, so in fact $\overline{I}(f) = \underline{I}(f)$, so f is integrable. $\square$

Here's an important application of this fact. Let S be a bounded region in the plane, and let Q be a rectangle that contains S. Let f be a bounded function on S. We extend f to a function $\widetilde{f}$ on Q by

$$\widetilde{f}(a) = \begin{cases} f(a) & \text{if } a \in S \\ 0 & \text{if } a \notin S \end{cases}.$$

If $\widetilde{f}$ is integrable on Q, then we say that f is integrable on S, and define

$$\iint_S f := \iint_Q \widetilde{f}.$$

Since we have extended the definition of f by setting it 0 outside of S, we see that actually this definition does not depend on the choice of Q (i.e., choosing a larger Q that contains S would not affect its integrability or the value of its integral).

By the previous theorem, it would suffice to know that the set where $\widetilde{f}$ fails to be continuous has content zero. For example, if f is continuous on S, then this roughly amounts to asking that the boundary of S has content zero.

A general class of such regions can be constructed from the following fact:

Lemma 3.12. *If $\varphi\colon [a,b] \to \mathbf{R}$ is a continuous function, then the graph of φ has content zero.*

Proof. This follows from Theorem 3.10 (it applies to intervals $[a,b]$ as a special case when $c = d$, for example). Pick $\varepsilon > 0$. Then we can break $[a,b]$ into finitely many subintervals so that the difference between the max and min of φ on these subintervals is at most $\varepsilon/(b-a)$. The graph of φ on such an interval is then contained in a rectangle of area ε times the length of the interval, so the whole graph of φ is contained in a finite union of rectangles whose total area is at most ε. $\square$

Given two functions $\varphi_1, \varphi_2\colon [a,b] \to \mathbf{R}$ such that $\varphi_1(x) \leq \varphi_2(x)$ for all $x \in [a,b]$, consider the region

$$S = \{(x,y) \mid a \leq x \leq b,\ \varphi_1(x) \leq y \leq \varphi_2(x)\}.$$

Apostol calls this a region of **Type I**. This is the area that is under the graph of φ_2 and above the graph of φ_1. The interior of S is defined by taking strict inequalities instead of weak ones in the definition of S. Similarly, we can define regions of **Type II** as those of the form

$$T = \{(x,y) \mid \varphi_1(y) \leq x \leq \varphi_2(y),\ a \leq y \leq b\}.$$

Theorem 3.13. *If S is a region of Type I, and $f\colon S \to \mathbf{R}$ is a function which is continuous on the interior of S, then $\iint_S f$ exists. Furthermore, we can compute it as an iterated integral:*

$$\iint_S f = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

Similarly, if S is a region of type II, and $f\colon S \to \mathbf{R}$ is a function which is continuous on the interior of S, then $\iint_S f$ exists. Furthermore, we can compute it as an iterated integral:

$$\iint_S f = \int_a^b \left(\int_{\varphi_1(y)}^{\varphi_2(y)} f(x,y)\,\mathrm{d}x \right) \mathrm{d}y.$$

Proof. We just discuss the Type I case since they are analogous.

First, note that φ_1, φ_2 are bounded functions since they are continuous on a finite interval. So S is a bounded region, and since f is continuous on it, f is also a bounded function. In particular, we can find a rectangle $Q = [a,b] \times [c,d]$ containing S, and extend f to a function $\widetilde{f}\colon Q \to \mathbf{R}$ by defining it to be 0 outside of S. By Lemma 3.12, the boundary of S has content 0, so by Theorem 3.11, $\widetilde{f}$ is integrable on Q, and so $\iint_S f = \iint_Q \widetilde{f}$ exists.

To prove the second part, we apply Theorem 3.4, or really the variant of it in Remark 3.5. Define $A(x) = \int_c^d \widetilde{f}(x,y)\,\mathrm{d}y$. Then $\widetilde{f}(x,y)$, as a function of y, is discontinuous in at most 2 places ($\varphi_1(x)$ and $\varphi_2(x)$), and hence $A(x)$ exists for each x. We can show that the function A is continuous as in the proof of Theorem 3.9, but there are more details to take care of, so we will omit it. In any case, then the integral of $A(x)$ exists, so we have

$$\iint_Q \widetilde{f} = \int_a^b \left(\int_c^d \widetilde{f}(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

Finally, we have $\int_c^d \widetilde{f}(x,y)\,\mathrm{d}y = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,\mathrm{d}y$ since $\widetilde{f}(x,y)$ is 0 outside of S. □

Example 3.14. Integrate the function $\sin(y^2)$ over the triangle T with vertices $\{(0,0),(1,0),(1,1)\}$.

The triangle is a Type II region with $\varphi_1(y) = 0$ and $\varphi_2(y) = y$ over the interval $[0,1]$. So we have

$$\begin{aligned}\iint_T \sin(y^2) &= \int_0^1 \left(\int_0^y \sin(y^2)\,\mathrm{d}x \right) \mathrm{d}y \\ &= \int_0^1 y\sin(y^2)\,\mathrm{d}y = -\frac{1}{2}\cos(y^2)|_0^1 = \frac{1}{2}(1-\cos(1)).\end{aligned}$$

The triangle is also a Type I region with $\varphi_1(x) = x$ and $\varphi_2(x) = 1$ over the interval $[0,1]$. So we have

$$\iint_T \sin(y^2) = \int_0^1 \left(\int_x^1 \sin(y^2)\,\mathrm{d}y \right) \mathrm{d}x.$$

However, this integral is more difficult to evaluate. □

We have already seen that the double integral can be interpreted as the volume of the region under the graph of a function on a rectangle. By the way we defined the double integral over a more general region S, when $\iint_S f$ exists, we may also interpret it as the (signed) volume of the region under the graph of f. In the special case $f(x,y) = 1$, the volume is just the area of the region S, so we see that

$$\iint_S 1 = \operatorname{area}(S).$$

This is not interesting for rectangles, and hence this is a genuinely new kind of computation in dimensions ≥ 2. If S is of Type I or Type II, this kind of computation isn't that surprising:

the formula in Theorem 3.13 gives

$$\operatorname{area}(S) = \int_a^b (\varphi_2(x) - \varphi_1(x))\,\mathrm{d}x$$

which you've already seen in 1-variable calculus. We'll see some ways to use this (and connect it to line integrals) for more general regions when we discuss Green's theorem.

Still, we can use the interpretation as volume to get some formulas. Here's a familiar example.

Example 3.15. Compute the volume of the sphere of radius r.

We can describe the top half of the sphere as the region underneath the function $f(x, y) = \sqrt{r^2 - x^2 - y^2}$ defined on the disk S of radius r. The disk S is a region of type I (and also type II). It can be described by the functions $\varphi_2(x) = \sqrt{r^2 - x^2}$ and $\varphi_1(x) = -\sqrt{r^2 - x^2}$ on the interval $[-r, r]$. So we can get the volume as the integral

$$2\int_{-r}^{r} \left(\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \sqrt{r^2 - x^2 - y^2}\,\mathrm{d}y \right) \mathrm{d}x.$$

We could evaluate this directly using trigonometric substitutions. Instead, note that the inner integral is computing the area of a half-disk of radius $\sqrt{r^2 - x^2}$, and hence the value is $\pi(r^2 - x^2)/2$. So we get

$$2\int_{-r}^{r} \frac{\pi}{2}(r^2 - x^2)\,\mathrm{d}x = \pi(2r^3 - \frac{2}{3}r^3) = \frac{4}{3}\pi r^3.$$

Similarly, we can compute the volumes of higher-dimensional spheres by iterating several integrals (using the higher dimensional analogues of the theorems we've discussed). For example, the volume of the 4-dimensional sphere of radius r is

$$2\int_{-r}^{r} \left(\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \left(\int_{-\sqrt{r^2-x^2-y^2}}^{\sqrt{r^2-x^2-y^2}} \sqrt{r^2 - x^2 - y^2 - z^2}\,\mathrm{d}z \right) \mathrm{d}y \right) \mathrm{d}x.$$

Again, the inner part is the integral we setup above for a 3-sphere of radius $\sqrt{r^2 - x^2}$, so this simplifies to

$$2\int_{-r}^{r} \frac{2}{3}\pi(r^2 - x^2)^{3/2}\,\mathrm{d}x = \frac{1}{2}\pi^2 r^4.$$

To get the final answer, one could use trigonometric substitutions. There might be better ways I'm not seeing. □

3.6. **Green's theorem.** In some cases, the integral of a function on a region can be reduced to a line integral along the boundary of the region. This is not entirely unfamiliar: think of the second fundamental theorem of calculus:

$$\int_a^b \varphi'(t)\,\mathrm{d}t = \varphi(b) - \varphi(a).$$

In this case, we are replacing a 1-dimensional integral by a "0-dimensional integral" along the boundary of the interval $[a, b]$ at the cost of having to take an antiderivative. Green's theorem says that something similar happens: 2-dimensional integrals can be replaced by 1-dimensional integrals. We will see other versions of this idea and time permitting, we'll discuss a complete generalization of this (general Stokes' theorem).

Let $\alpha\colon [a,b] \to \mathbf{R}^2$ be a continuous planar curve. Recall that α is **closed** if $\alpha(a) = \alpha(b)$. We say that α is **simple** if α is injective on $[a,b)$, i.e., if $a \le t_0 < t_1 < b$, then $\alpha(t_0) \ne \alpha(t_1)$. This just means that α does not cross itself. We say that α is a **Jordan curve** if it is both simple and closed.

The *Jordan curve theorem* tells us that if we remove the image of α from $\mathbf{R}^2$, we are left with 2 connected regions: a bounded region, which we call the **interior**, and an unbounded region, which we call the **exterior**.

Finally, given a Jordan curve parametrized by α, we can say whether it is traveling **clockwise** or **counterclockwise**. We don't have a good language to define this rigorously (see the section in Apostol on winding numbers for a way to do this), so we will rely on the intuitive notion.

Example 3.16. If $\alpha\colon [0,1] \to \mathbf{R}^2$ is given by $\alpha(t) = (\cos(t), \sin(t))$, then the image is the unit circle and it is traveling counterclockwise. The interior is $\{(x,y) \mid x^2 + y^2 < 1\}$ and the exterior is $\{(x,y) \mid x^2 + y^2 > 1\}$. $\square$

Theorem 3.17 (Green's theorem)**.** *Let C be a piecewise smooth Jordan curve parametrized by α in a counterclockwise direction, and let R be the interior of C union with C. Let S be an open set that contains R, and let $f = (f_1, f_2)\colon S \to \mathbf{R}^2$ be a continuously differentiable vector field. Then*

$$\iint_R \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) = \int f \cdot \mathrm{d}\alpha.$$

Remark 3.18. Much of what we have said up to this point is valid for an arbitrary number of dimensions, but it is important to keep in mind that Green's theorem is strictly about 2-dimensional integrals! We will see later how it generalizes, but right now it is not clear. $\square$

Remark 3.19. I've written the theorem in a way that reflects our previous notation, but it is usually written with different notation: $P = f_1$ and $Q = f_2$, and the right side of the formula is usually written

$$\ointctrclockwise_C (P\,\mathrm{d}x + Q\,\mathrm{d}y).$$

The integral symbol denotes that we are traveling around C in the counterclockwise direction via a parametrization α, and the inner integral is an alternate notation for

$$(P(\alpha_1(t), \alpha_2(t))\alpha_1'(t) + Q(\alpha_1(t), \alpha_2(t))\alpha_2'(t))\,\mathrm{d}t. \qquad \square$$

We will go through the proof of Green's theorem in some easy cases later, but first, let's see how this can be applied to compute areas. Given a bounded region S whose boundary has content 0, we have

$$\mathrm{area}(S) = \iint_S 1.$$

Assuming further that the boundary is a piecewise smooth Jordan curve C, we can apply Green's theorem to get

$$\mathrm{area}(S) = \ointctrclockwise_C (P\,\mathrm{d}x + Q\,\mathrm{d}y)$$

for any choice of (P,Q) such that $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$. We can take $(P,Q) = ((c-1)y, cx)$ for any real number c, so we get the following consequence of Green's theorem:

Theorem 3.20. *Let S be a bounded region whose boundary is a piecewise smooth Jordan curve C which is parametrized by $\alpha\colon [a,b] \to \mathbf{R}^2$. Write the components of α as $\alpha(t) = (X(t), Y(t))$. Then for any real number c, we have*

$$\operatorname{area}(S) = \int_a^b ((c-1)X'(t)Y(t) + cX(t)Y'(t))\,\mathrm{d}t.$$

If we choose $c = 0$ or $c = 1$, the formula above simplifies a lot. However, the next example shows that different choices might give easier integrals.

Example 3.21. Let S be the unit disk. Then C is the unit circle and $\alpha(t) = (\cos(t), \sin(t))$ on $[0, 2\pi]$. If we choose $c = 1$, then we get

$$\int_0^{2\pi} \cos^2(t)\,\mathrm{d}t,$$

which can be solved using a double-angle identity. Alternatively, we can choose $c = 1/2$ and then we get

$$\frac{1}{2}\int_0^{2\pi} (\sin^2(t) + \cos^2(t))\,\mathrm{d}t = \frac{1}{2}\int_0^{2\pi} \mathrm{d}t = \pi. \qquad \square$$

Example 3.22. This example illustrates a piecewise defined curve (even though the end result is easy). Consider the triangle with vertices $(0,0), (1,0), (0,1)$. We can parametrize it by $\alpha\colon [0,3] \to \mathbf{R}^2$ with $\alpha(t) = (X(t), Y(t))$ and

$$X(t) = \begin{cases} t & \text{if } 0 \le t \le 1 \\ 2-t & \text{if } 1 \le t \le 2 \\ 0 & \text{if } 2 \le t \le 3 \end{cases}, \qquad Y(t) = \begin{cases} 0 & \text{if } 0 \le t \le 1 \\ t-1 & \text{if } 1 \le t \le 2 \\ 3-t & \text{if } 2 \le t \le 3 \end{cases}.$$

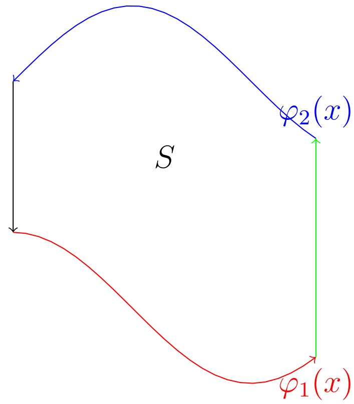
We'll take $c = 1$ in the theorem above. Then

$$X(t)Y'(t) = \begin{cases} 0 & \text{if } 0 \le t \le 1 \\ 2-t & \text{if } 1 \le t \le 2 \\ 0 & \text{if } 2 \le t \le 3 \end{cases},$$

so the area of the triangle is given by

$$\int_1^2 (2-t)\mathrm{d}t = 2t - \frac{t^2}{2}\Big|_1^2 = 2 - \frac{3}{2} = \frac{1}{2}. \qquad \square$$

3.7. **Proof of Green's theorem.** We will only prove Green's theorem in some special cases. First we start with the case when S is a region of Type I and $Q = 0$. Consider the picture below:



Since $Q = 0$, we're trying to show that

$$-\iint_S \frac{\partial P}{\partial y} = \int (P, 0) \cdot \mathrm{d}\alpha$$

where α parametrizes the boundary of S. First let's simplify the left hand side. We can write it as an iterated integral:

$$-\iint_R \frac{\partial P}{\partial y} = -\int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} \frac{\partial P}{\partial y} \,\mathrm{d}y \right) \mathrm{d}x = -\int_a^b (P(x, \varphi_2(x)) - P(x, \varphi_1(x)) \,\mathrm{d}x.$$

For the right hand side, we use the parametrization $\alpha(t) = (t, \varphi_1(t))$ from $[a, b]$ for the bottom part and the parametrization $\gamma(t) = (t, \varphi_2(t))$ from $[a, b]$ for the top part (but notice we're going in the wrong direction here). For the right side, we use $\beta(t) = (b, t)$ from $[\varphi_1(b), \varphi_2(b)]$ and for the left side, we use $\delta(t) = (a, t)$ from $[\varphi_1(a), \varphi_2(a)]$ (again, we're going in the wrong direction here). Then the right hand side becomes

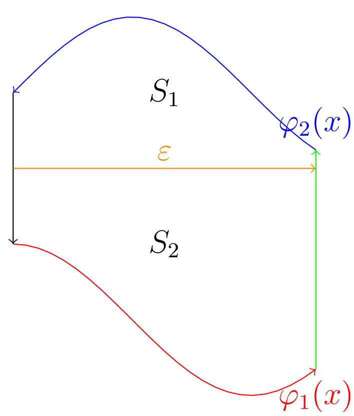
$$\int (P, 0) \cdot \mathrm{d}\alpha + \int (P, 0) \cdot \mathrm{d}\beta - \int (P, 0) \cdot \mathrm{d}\gamma - \int (P, 0) \cdot \mathrm{d}\delta = \int_a^b P(t, \varphi_1(t)) \,\mathrm{d}t - \int_a^b P(t, \varphi_2(t)) \,\mathrm{d}t,$$

which is what we wanted.

In a similar way, we can show that Green's theorem holds if S is a Type II region and $P = 0$.

By adding together these two cases, **we have shown that Green's theorem holds for regions S that are both Type I and Type II.**

To do the general case, we can chop up our region into smaller pieces. Consider the following:

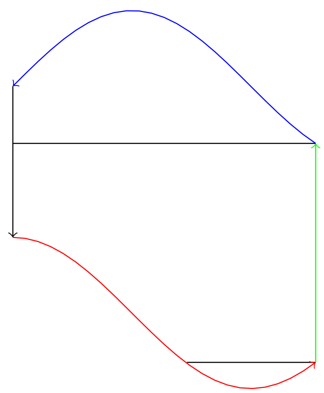


Let ∂S denote the boundary of a region S. Then we have:

$$\iint_S = \iint_{S_1} + \iint_{S_2}, \qquad \oint_{\partial S} = \oint_{\partial S_1} + \oint_{\partial S_2}.$$

The second identity follows because ∂S_1 uses ε while ∂S_2 uses ε in the opposite direction, so that these terms cancel when we sum them together. Hence, if we know Green's theorem holds for the pieces S_1 and S_2, then we also know that it holds for S. We get the following: **Green's theorem holds for a region S if we can decompose it into finitely many pieces which are both of Type I and Type II.**

Fortunately, many regions we might think of have such a decomposition. Going beyond to things that don't, we'd need a more general proof, but we will stop here. In our example, we might do this:



3.8. **Change of variables in a double integral.** Our aim now is to give a version of "u-substitution" for double integrals. Recall that for 1-variable integrals, we have identities of the form

$$\int_a^b f(g(t))g'(t)\,\mathrm{d}t = \int_{g(a)}^{g(b)} f(x)\,\mathrm{d}x$$

which we think of as being related by the substitution $x = g(t)$. Here the assumption is that g' and f are continuous.

In the 2-variable case, a change of variables should involve 2 functions

$$x = X(u, v), \qquad y = Y(u, v).$$

We will assume that the partial derivatives of X and Y exist and are continuous. So we can form the **Jacobian determinant**

$$J(u, v) := \det \begin{bmatrix} \dfrac{\partial X}{\partial u} & \dfrac{\partial X}{\partial v} \\ \dfrac{\partial Y}{\partial u} & \dfrac{\partial Y}{\partial v} \end{bmatrix}.$$

Given a region T (thought of as living in the u, v-plane), we can define a new region

$$S = \{(X(u, v), Y(u, v)) \mid (u, v) \in T\}$$

(thought of as living in the x, y-plane). We will assume that there is a subset of T of content 0 so that the mapping from the complement of this subset to S is injective, i.e., if $(X(u, v), Y(u, v)) = (X(u', v'), Y(u', v'))$ and (u, v), then $(u, v) = (u', v')$. Furthermore, we will assume that $J(u, v) \geq 0$ for all $(u, v) \in T$ and that the set of (u, v) where $J(u, v) = 0$ has content 0.

The analogous change of variables formula (which we prove later) is:

Theorem 3.23. *With the notation as above, we have*

$$\iint_T f(X(u,v),Y(u,v))J(u,v)\,\mathrm{d}u\mathrm{d}v = \iint_S f(x,y)\,\mathrm{d}x\mathrm{d}y.$$

We haven't been writing the $\mathrm{d}x\mathrm{d}y$ for multiple integrals before, but we do it here to emphasize that we are changing the variables.

Remark 3.24. Here is a heuristic for the formula (hopefully to be made precise later). We can write

$$\mathrm{d}x = \frac{\partial X}{\partial u}\mathrm{d}u + \frac{\partial X}{\partial v}\mathrm{d}v, \qquad \mathrm{d}y = \frac{\partial Y}{\partial u}\mathrm{d}u + \frac{\partial Y}{\partial v}\mathrm{d}v.$$

If we expand out the product $\mathrm{d}x\mathrm{d}y$, then we get $J(u,v)\,\mathrm{d}u\mathrm{d}v$ if we adopt the rules that $\mathrm{d}u\mathrm{d}u = 0 = \mathrm{d}v\mathrm{d}v$ and $\mathrm{d}u\mathrm{d}v = -\mathrm{d}v\mathrm{d}u$. At this point it doesn't make much sense to do this, though it is a convenient way to remember the formula (the one thing we need to remember is that the symbols $\mathrm{d}x$ pick up a sign whenever we move them past each other). □

An important example is given by **polar coordinates**. In this case, it is traditional to use r (radius) and θ (angle) instead of u and v. The change of variables is

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

Instead of labeling points by their (x,y)-coordinate, polar coordinates specify points by two pieces of information: the angle (measured in radians) θ of the vector from the origin to that point and its distance from the origin. To make our assumptions above satisfied, we will assume that T lies in the region with $r \ge 0$ and $0 \le \theta < 2\pi$.

In this case, our Jacobian determinant becomes

$$J(r,\theta) := \det \begin{bmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{bmatrix} = r\cos^2\theta + r\sin^2\theta = r \ge 0,$$

so the change of variables formula becomes

$$\iint_S f(x,y)\,\mathrm{d}x\mathrm{d}y = \iint_T f(r\cos\theta, r\sin\theta)r\,\mathrm{d}r\mathrm{d}\theta.$$

Example 3.25. The point of polar coordinates is that rectangles in the r,θ-plane become disks in the x,y-plane. For example, let T be the rectangle with $0 \le r \le a$ and $0 \le \theta \le 2\pi$ let $f(x,y) = 1$. Then S is the disk of radius a, and we have

$$\iint_S 1\,\mathrm{d}x\mathrm{d}y = \iint_T r\,\mathrm{d}r\mathrm{d}\theta = \int_0^{2\pi}\int_0^a r\,\mathrm{d}r\mathrm{d}\theta = \int_0^{2\pi} \frac{a^2}{2}\,\mathrm{d}\theta = \pi a^2.$$

This is nothing new, but the point is that the integral in polar coordinates is much simpler than the integral in Cartesian coordinates. □

Example 3.26. Here's a more striking example where we can compute something we couldn't do before. Consider the problem of evaluating the improper integral

$$I = \int_0^\infty e^{-x^2}\,\mathrm{d}x.$$

Techniques from first-year calculus won't help much since the antiderivative of e^{-x^2} isn't any familiar function, though you can use first-year calculus to show that I is a finite value.

It turns out it is much easier to evaluate I^2. We haven't dealt with infinite regions, so what follows is missing justification since I just want to emphasize the idea rather than the details. We can compute this as an iterated integral:

$$\int_0^\infty \left(\int_0^\infty e^{-x^2-y^2}\,\mathrm{d}x\right)\mathrm{d}y = \int_0^\infty e^{-y^2}\left(\int_0^\infty e^{-x^2}\,\mathrm{d}x\right)\mathrm{d}y = \left(\int_0^\infty e^{-x^2}\,\mathrm{d}x\right)\left(\int_0^\infty e^{-y^2}\,\mathrm{d}y\right) = I^2.$$

Here we are integrating over the first quadrant. In polar coordinates, this is the region $0 \le r \le \infty$ and $0 \le \theta \le \pi/2$. Hence we also get

$$I^2 = \int_0^{\pi/2}\left(\int_0^\infty e^{-r^2} r\,\mathrm{d}r\right)\mathrm{d}\theta = \int_0^{\pi/2}\left(-\frac{1}{2}e^{-r^2}\Big|_0^\infty\right)\mathrm{d}\theta = \int_0^{\pi/2}\frac{1}{2}\,\mathrm{d}\theta = \frac{\pi}{4}.$$

Since $I \ge 0$ (because $e^{-x^2} \ge 0$), we conclude that

$$\int_0^\infty e^{-x^2}\,\mathrm{d}x = \frac{\sqrt{\pi}}{2}. \qquad \square$$

3.9. **Proof of change of variables formula.** Recall that the integral of a function is defined in terms of the integral of step functions, which in turn relied on the area of a rectangle. The outline of the proof for the change of variables formula

$$\iint_T f(X(u,v), Y(u,v))J(u,v)\,\mathrm{d}u\mathrm{d}v = \iint_S f(x,y)\,\mathrm{d}x\mathrm{d}y$$

follows this:

- First we show the formula holds when S is a rectangle and $f(x,y) = 1$.
- Next we show the formula holds when S is a rectangle and f is a step function.
- Finally, we move to a general integrable function f.

To avoid complications, we make the following additional assumptions:

- The Jacobian $J(u,v)$ is never 0. Note that this does not hold even in the example of polar coordinates (since $r = 0$ is a possibility). It can be dealt with, but we won't do it here.
- The second derivatives of X and Y exist and are continuous.
- The mapping (X, Y) is injective everywhere, and has an inverse.

3.9.1. *The case when $f(x,y) = 1$.* The right side of the formula is $\iint_S 1\,\mathrm{d}x\mathrm{d}y$ where S is a rectangle. By Green's theorem applied to the vector field $(0, x)$, we have

$$\iint_S 1\,\mathrm{d}x\mathrm{d}y = \oint_{\partial S} x\,\mathrm{d}y$$

where ∂S is the boundary of S. For the left side of the formula, we have

$$J(u,v) = \frac{\partial X}{\partial u}\frac{\partial Y}{\partial v} - \frac{\partial X}{\partial v}\frac{\partial Y}{\partial u} = \frac{\partial}{\partial u}\left(X\frac{\partial Y}{\partial v}\right) - \frac{\partial}{\partial v}\left(X\frac{\partial Y}{\partial u}\right).$$

By another application of Green's theorem to the vector field $\left(X\frac{\partial Y}{\partial u}, X\frac{\partial Y}{\partial v}\right)$, we have

$$\iint_T J(u,v)\,\mathrm{d}u\mathrm{d}v = \oint_{\partial T}\left(X\frac{\partial Y}{\partial u}\mathrm{d}u + X\frac{\partial Y}{\partial v}\mathrm{d}v\right)$$

where ∂T is the boundary of T. So to finish this case, we want to show that

$$\oint_{\partial T}\left(X\frac{\partial Y}{\partial u}\mathrm{d}u + X\frac{\partial Y}{\partial v}\mathrm{d}v\right) = \oint_{\partial S} x\,\mathrm{d}y.$$

Let $\alpha\colon [a,b] \to \mathbf{R}^2$ be a counterclockwise parametrization of ∂T. Write $\alpha(t) = (U(t), V(t))$. Then we have

$$\iint_T J(u,v)\,\mathrm{d}u\mathrm{d}v = \oint_{\partial T} \left(X\frac{\partial Y}{\partial u}\mathrm{d}u + X\frac{\partial Y}{\partial v}\mathrm{d}v \right) = \int_a^b X(U(t),V(t)) \left(\frac{\partial Y}{\partial u}U'(t) + \frac{\partial Y}{\partial v}V'(t) \right) \mathrm{d}t$$

Define $\beta\colon [a,b] \to \mathbf{R}^2$ by

$$\beta(t) = (X(U(t),V(t)), Y(U(t),V(t)))$$

The derivative is

$$\beta'(t) = \left(\frac{\partial X}{\partial u}U'(t) + \frac{\partial X}{\partial v}V'(t), \frac{\partial Y}{\partial u}U'(t) + \frac{\partial Y}{\partial v}V'(t) \right)$$

So the last expression above is $\int (X,0)\cdot \mathrm{d}\beta$. Finally, β gives a parametrization for ∂S, so we get

$$\oint_{\partial T} \left(X\frac{\partial Y}{\partial u}\mathrm{d}u + X\frac{\partial Y}{\partial v}\mathrm{d}v \right) = \pm \oint_{\partial S} x\,\mathrm{d}y$$

where the sign is $+$ if β is counterclockwise, and $-$ otherwise. The left integral is the same as $\iint_T J(u,v)\,\mathrm{d}u\mathrm{d}v$, which is positive since $J(u,v)$ is positive. The right integral is also positive since it is the same as $\iint_S 1\mathrm{d}x\mathrm{d}y$, so we conclude that the sign is $+$.

3.9.2. *The case when $f(x,y)$ is a step function.* If f is a step function on S, then partition S into subrectangles $S_1,\dots,S_n$ so that f is constant on each. Let c_i be the value of f on S_i. Since (X,Y) is invertible, let T_i be the region in the u,v plane that maps to S_i. Then by the previous case, we have

$$\iint_S f(x,y) = \sum_{i=1}^n c_i \iint_{S_i} 1\,\mathrm{d}x\mathrm{d}y = \sum_{i=1}^n c_i \iint_{T_i} J(u,v)\,\mathrm{d}u\mathrm{d}v = \iint_T f(X(u,v),Y(u,v))J(u,v)\,\mathrm{d}u\mathrm{d}v,$$

where the last equality uses that $\iint_T = \sum_{i=1}^n \iint_{T_i}$.

3.9.3. *The case of a general integrable function f.* Now suppose f is an integrable function on R. Pick step functions s,t such that $s(x,y) \le f(x,y) \le t(x,y)$ for all x,y. In particular, we also have

$$s(X(u,v),Y(u,v)) \le f(X(u,v),Y(u,v)) \le t(X(u,v),Y(u,v))$$

for all u,v since $X(u,v),Y(u,v)$ is just a particular instance of x,y. Since $J(u,v) > 0$, we can multiply this inequality by $J(u,v)$ and integrate over T to get

$$\begin{aligned} \iint_T s(X(u,v),Y(u,v))J(u,v)\mathrm{d}u\mathrm{d}v &\le \iint_T f(X(u,v),Y(u,v))J(u,v)\mathrm{d}u\mathrm{d}v \\ &\le \iint_T t(X(u,v),Y(u,v))J(u,v)\mathrm{d}u\mathrm{d}v. \end{aligned}$$

By the case we just saw, this can be rewritten as

$$\iint_S s(x,y)\mathrm{d}x\mathrm{d}y \le \iint_T f(X(u,v),Y(u,v))J(u,v)\mathrm{d}u\mathrm{d}v \le \iint_S t(x,y)\mathrm{d}x\mathrm{d}y.$$

Since this is true for all s,t with $s \le f \le t$, we conclude that the middle integral satisfies

$$\underline{I}(f) \le \iint_T f(X(u,v),Y(u,v))J(u,v)\mathrm{d}u\mathrm{d}v \le \overline{I}(f).$$

But f is integrable, so the outer two values are the same, from which we conclude that

$$\iint_T f(X(u,v),Y(u,v))J(u,v)\mathrm{d}u\mathrm{d}v = \int_S f(x,y)\mathrm{d}x\mathrm{d}y.$$

3.10. **More than 2 dimensions.** As I mentioned before, most everything in this section can be adapted to n-dimensions for any n. Generalizations of Green's theorem will be discussed later. But here are some highlights of what goes through in a straightforward manner.

3.10.1. *Type I/II regions.* Given a region $Q \subseteq \mathbf{R}^2$ and functions $\varphi_1, \varphi_2 \colon Q \to \mathbf{R}$, we can define an analogue of Type I/II regions to be those of the form

$$S = \{(x,y,z) \mid (x,y) \in Q,\ \varphi_1(x,y) \le z \le \varphi_2(x,y)\}.$$

If f is a continuous function on S, we have

$$\iiint_S f = \iint_Q \int_{\varphi_1(x,y)}^{\varphi_2(x,y)} f(x,y,z)\,\mathrm{d}z.$$

We can also swap z with either x and y and get different kinds of regions.

3.10.2. *Change of variables.* Suppose we have variables $u_1,\dots,u_n$ and $x_1,\dots,x_n$ related by the equations

$$x_i = X_i(u_1,\dots,u_n) = X_i(\mathbf{u}).$$

where the third expression is shorthand for the second one. We can form the Jacobian determinant

$$J(\mathbf{u}) = \det \begin{bmatrix} \frac{\partial X_1}{\partial u_1} & \cdots & \frac{\partial X_1}{\partial u_n} \\ \vdots & & \vdots \\ \frac{\partial X_n}{\partial u_1} & \cdots & \frac{\partial X_n}{\partial u_n} \end{bmatrix}.$$

Again, assuming $(X_1,\dots,X_n)$ gives an injective mapping and the Jacobian is never 0 (or some slight relaxation of this), we have a change of variables formula for integrals:

$$\iiint_S f(x_1,\dots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n = \iiint_T f(X_1(\mathbf{u}),\dots,X_n(\mathbf{u}))J(\mathbf{u})\,\mathrm{d}u_1\cdots\mathrm{d}u_n$$

where T is some region in $u_1,\dots,u_n$-space, and

$$S = \{(X_1(\mathbf{u}),\dots,X_n(\mathbf{u})) \mid (u_1,\dots,u_n) \in T\}.$$

Here are 2 important examples when $n=3$.

Example 3.27 (Cylindrical coordinates)**.** Cyclindrical coordinates are essentially the same as polar coordinates. Instead of u_1,u_2,u_3, we write r,θ,z. Here, given a point (x,y,z), we are keeping z the same and (r,θ) are the radius and angle of (x,y) in the x,y-plane. Specifically, we have

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z$$

where we stick to the region where $r \ge 0$ and $0 \le \theta < 2\pi$. The Jacobian determinant is

$$\det \begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = r\cos^2\theta + r\sin^2\theta = r \ge 0.$$

The point here is that cylinders are "rectangles" in the r, θ, z-variables. If we have a cylinder S whose base circle is centered at $(0,0)$ and has radius R and whose height goes from h_1 to h_2, then integrating a continuous function f over S becomes

$$\iiint_S f = \int_{h_1}^{h_2} \left(\int_0^{2\pi} \left(\int_0^R f(r\cos\theta, r\sin\theta, z) r \,\mathrm{d}r \right) \mathrm{d}\theta \right) \mathrm{d}z.$$

Since the bounds of the integrals don't depend on each other, you can also exchange the order of integration if needed. □

Example 3.28 (Spherical coordinates)**.** In spherical coordinates, we represent a point in 3-space by its distance ρ from the origin, the angle φ between the positive z-axis and the vector from $(0,0,0)$ to that point, and the angle θ used in cyclindrical coordinates. Since the length of (x,y) is given by $\rho \sin\varphi$, we have

$$x = \rho\cos\theta\sin\varphi, \qquad y = \rho\sin\theta\sin\varphi, \qquad z = \rho\cos\varphi.$$

We also have the restrictions $\rho \ge 0$, $0 \le \theta < 2\pi$, and $0 \le \varphi \le \pi$. The Jacobian determinant is (expanding along the last row)

$$\begin{aligned} J(\rho,\varphi,\theta) &= \det \begin{bmatrix} \cos\theta\sin\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & \rho\sin\theta\cos\varphi & \rho\cos\theta\sin\varphi \\ \cos\varphi & -\rho\sin\varphi & 0 \end{bmatrix} \\ &= \cos\varphi(\rho^2\cos^2\theta\cos\varphi\sin\varphi + \rho^2\sin^2\theta\cos\varphi\sin\varphi) \\ &\qquad + \rho\sin\varphi(\rho\cos^2\theta\sin^2\varphi + \rho\sin^2\theta\sin^2\varphi) \\ &= \rho^2\cos^2\varphi\sin\varphi + \rho^2\sin^3\varphi = \rho^2\sin\varphi. \end{aligned}$$

The last quantity is ≥ 0 since $0 \le \varphi \le \pi$.

The point here is that a 3-dimensional sphere S centered at $(0,0,0)$ is a "rectangle" in the ρ, φ, θ-variables. Namely, if the radius is R, then we take $0 \le \rho \le R$, $0 \le \theta < 2\pi$, and $0 \le \varphi \le \pi$, so integrating a continuous fnuction f over S becomes

$$\iiint_S f = \int_0^{2\pi} \left(\int_0^{\pi} \left(\int_0^R f(\rho\cos\theta\sin\varphi, \rho\sin\theta\sin\varphi, \rho\cos\varphi)\rho^2\sin\varphi \,\mathrm{d}\rho \right) \mathrm{d}\varphi \right) \mathrm{d}\theta.$$

Again, you can exchange the order of integration if needed. □

4. Surface integrals

Surface integrals are 2-dimensional analogues of line integrals. To line up notation: the surface integrals defined by Apostol are analogous to line integrals with respect to arc length. So surface integrals are defined for scalar fields, and we'll see what the analogue for vector fields is. One thing surface integrals will do for us is to give us a 3-dimensional version of Green's theorem (also called the divergence theorem or Gauss' theorem): the integrals of certain functions over 3-dimensional regions can be rewritten as a surface integral over its boundary.

4.1. **Parametrizations of surfaces.** For line integrals, we represented curves in space as the image of an interval $\alpha\colon [a,b] \to \mathbf{R}^n$. For this section, we'll do the same for surfaces, but stick to $\mathbf{R}^3$. Since we want something 2-dimensional, the interval $[a,b]$ gets replaced by a 2-dimensional connected set $T \subseteq \mathbf{R}^2$ and α becomes a function $\mathbf{r}\colon T \to \mathbf{R}^3$ (the change in notation is to emphasize this is a surface and not a curve). We usually write the component functions of $\mathbf{r}$ as X, Y, Z and the variables in T as u, v. So our surface S is the set

$$S = \{(X(u,v), Y(u,v), Z(u,v)) \mid (u,v) \in T\}.$$

The function $\mathbf{r} = (X, Y, Z)$ gives a **parametric representation** of S.

Example 4.1. The 2-dimensional sphere of radius r has an explicit representation given by the function $F(x,y,z) = x^2 + y^2 + z^2 - r^2$. We can also parametrize it by taking $T = \{(u,v) \mid 0 \le u \le 2\pi, \quad -\pi/2 \le v \le \pi/2\}$ and

$$X(u,v) = r\cos u \cos v, \qquad Y(u,v) = r \sin u \cos v, \qquad Z(u,v) = r \sin v. \qquad \square$$

Example 4.2. We can get a cone of height h with vertex at the origin whose radius at height r is given by r with the following parametrization:

$$X(u,v) = v\cos u, \qquad Y(u,v) = v \sin u, \qquad Z(u,v) = v.$$

Here the domain is $0 \le v \le h$ and $0 \le u \le 2\pi$. To get a cone with a different ratio, say we want the radius at height r to be cr for some constant c, then we take

$$X(u,v) = cv\cos u, \qquad Y(u,v) = cv \sin u, \qquad Z(u,v) = v.$$

By making $c > 1$, we get wider cones than the first example, and $0 < c < 1$ gives us skinnier cones. $\square$

Example 4.3. If a surface S is represented as the graph of a 2-variable function, i.e., we have a function $f(x,y)$ so that S it is the set of points $\{(x, y, f(x,y)) \mid (x,y) \in T\}$ for some region T, then we call this an **explicit representation** of S and we can parametrize it by $\mathbf{r}\colon T \to S$ using

$$X(u,v) = u, \qquad Y(u,v) = v, \qquad Z(u,v) = f(u,v).$$

For example, the top hemisphere of the sphere of radius R has an explicit representation given by the function $f(x,y) = \sqrt{R^2 - x^2 - y^2}$ over the region T which is the disk of radius R. While this only gives half of the sphere, when we are doing surface integrals, we can always chop up our surface into different pieces and integrate on each separately. $\square$

Example 4.4. Surfaces obtained by rotating the graph of a one-variable function around one of the axes also have an easy parametrization. If $f(x)$ is the function, and we want to rotate the graph of f around the x-axis, we can take

$$X(u,v) = u, \qquad Y(u,v) = f(u)\cos v, \qquad Z(u,v) = f(u) \sin v.$$

where here u ranges over the domain of f (which we assume is a finite interval if want a bounded region) and $0 \le v \le 2\pi$. $\square$

Some terminology and notation: we will denote S by $\mathbf{r}(T)$ to highlight the function $\mathbf{r}$ and its domain, and call $\mathbf{r}(T)$ a **parametric surface**. We will usually assume that $\mathbf{r}$ is continuous. If $\mathbf{r}$ is injective, i.e., $\mathbf{r}(u,v) = \mathbf{r}(u',v')$ if and only if $(u,v) = (u',v')$, then $\mathbf{r}(T)$ is a **simple parametric surface**.

4.2. The fundamental vector product.

Definition 4.5. Given two 3-dimensional vectors $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ (the entries are either real numbers or functions), their **cross product** is another 3-dimensional vector $\mathbf{a} \times \mathbf{b}$ defined by

$$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2, -a_1b_3 + a_3b_1, a_1b_2 - a_2b_1).$$

Here is a way to remember the formula: if we write vectors as $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are the standard basis vectors for $\mathbf{R}^3$, then

$$\mathbf{a} \times \mathbf{b} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}. \qquad \square$$

The cross product is special to 3-dimensions. Here's a few basic properties of the cross product:

- $\mathbf{a} \times \mathbf{a} = \mathbf{0}$
- $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
- $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$
- $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$
- $\mathbf{a} \times \mathbf{b}$ is orthogonal to $\mathbf{a}$ and $\mathbf{b}$, i.e., $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = 0$ and $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} = 0$
- $\|\mathbf{a} \times \mathbf{b}\|$ is the area of the parallelogram formed by $\mathbf{a}$ and $\mathbf{b}$.

Let $\mathbf{r}(T)$ be a parametric surface. We now assume that $\mathbf{r}$ is differentiable and define

$$\frac{\partial \mathbf{r}}{\partial u} = \left(\frac{\partial X}{\partial u}, \frac{\partial Y}{\partial u}, \frac{\partial Z}{\partial u}\right), \qquad \frac{\partial \mathbf{r}}{\partial v} = \left(\frac{\partial X}{\partial v}, \frac{\partial Y}{\partial v}, \frac{\partial Z}{\partial v}\right).$$

Given two functions F, G in variables x, y, we'll use the shorthand:

$$\frac{\partial(F,G)}{\partial(x,y)} = \det \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial G}{\partial x} \\ \frac{\partial F}{\partial y} & \frac{\partial G}{\partial y} \end{pmatrix}$$

Note that $\frac{\partial(F,G)}{\partial(x,y)} = -\frac{\partial(G,F)}{\partial(x,y)}$.

The **fundamental vector product** is

$$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = \frac{\partial(Y,Z)}{\partial(u,v)}\mathbf{i} + \frac{\partial(Z,X)}{\partial(u,v)}\mathbf{j} + \frac{\partial(X,Y)}{\partial(u,v)}\mathbf{k}.$$

A point $\mathbf{r}(a,b) \in \mathbf{r}(T)$ is a **regular point** of $\mathbf{r}$ if both $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ are continuous at (a,b) and the fundamental vector product is nonzero at (a,b). All other points are called **singular points** of $\mathbf{r}$. The parametric surface $\mathbf{r}(T)$ is **smooth** if all points of T are regular. Warning: this depends on $\mathbf{r}$, and not just the image of T, meaning that there are surfaces that have both smooth parametrizations and also non-smooth parametrizations.

Theorem 4.6. *Let $\mathbf{r}(T)$ be a smooth parametric surface, let D be a smooth curve in T, and let $C = \mathbf{r}(D)$ be the image of C under $\mathbf{r}$. Then for each $(a,b) \in D$, the fundamental vector product $\frac{\partial \mathbf{r}}{\partial u}(a,b) \times \frac{\partial \mathbf{r}}{\partial v}(a,b)$ is orthogonal to the tangent vector of C at $\mathbf{r}(a,b)$.*

Proof. Let $\alpha(t) = (U(t), V(t))$ be a parametrization for D. We get a parametrization ρ for C by composing with $\mathbf{r}$:

$$\rho(t) = \mathbf{r}(U(t), V(t)) = (X(U(t), V(t)), Y(U(t), V(t)), Z(U(t), V(t)))$$

with derivative

$$\begin{aligned}\rho'(t) &= (\nabla X(U(t),V(t))\cdot\alpha'(t), \nabla Y(U(t),V(t))\cdot\alpha'(t), \nabla Z(U(t),V(t))\cdot\alpha'(t)) \\ &= \frac{\partial \mathbf{r}}{\partial u}(U(t),V(t))U'(t) + \frac{\partial \mathbf{r}}{\partial v}(U(t),V(t))V'(t)\end{aligned}$$

We know that $\frac{\partial \mathbf{r}}{\partial u}(U(t),V(t)) \times \frac{\partial \mathbf{r}}{\partial v}(U(t),V(t))$ is orthogonal to both $\frac{\partial \mathbf{r}}{\partial u}(U(t),V(t))$ and $\frac{\partial \mathbf{r}}{\partial v}(U(t),V(t))$, so it is also orthogonal to $\rho'(t)$ by linearity of cross products. $\square$

Since the fundamental vector product at (a,b) is orthogonal to the tangent vector of all curves that pass through $\mathbf{r}(a,b)$, we see that it is actually normal to the tangent plane at $\mathbf{r}(a,b)$. Since we have a normal vector and a point on the tangent plane, we can write down an equation for it.

Recall that if $\mathbf{n} = (n_1,n_2,n_3)$ is normal to a plane, and (a,b,c) is a point on that plane, then that plane is the set of (x,y,z) satisfying the equation

$$n_1(x-a) + n_2(y-b) + n_3(z-c) = 0.$$

4.3. **Definition of the surface integral.** Consider the parametrization $\mathbf{r}\colon T \to S$. From now on, we'll assume that $\mathbf{r}$ is differentiable. Consider a rectangle in T with sides Δu and Δv (these are meant to indicate small values). Under $\mathbf{r}$, this turns into a region which we can linearly approximate using the partial derivatives $\frac{\partial \mathbf{r}}{\partial u}\Delta u$ and $\frac{\partial \mathbf{r}}{\partial v}\Delta v$. By properties of cross product mentioned earlier, the area of this approximation is $\|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\|\Delta u \Delta v$. Similar to how one gets the arc length for a parametrized curve, we can write down a formula for the surface area $a(S)$ of S by covering T by smaller and smaller rectangles and summing the result:

$$a(S) = \iint_T \left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| \,\mathrm{d}u\mathrm{d}v.$$

Example 4.7. Consider the parametrization of the sphere S of radius r given in Example 4.1:

$$X(u,v) = r\cos u\cos v, \qquad Y(u,v) = r\sin u\cos v, \qquad Z(u,v) = r\sin v$$

where T is the region with $0 \le u \le 2\pi$ and $-\pi/2 \le v \le \pi/2$. We compute the partials:

$$\frac{\partial \mathbf{r}}{\partial u} = (-r\sin u\cos v, r\cos u\cos v, 0), \qquad \frac{\partial \mathbf{r}}{\partial v} = (-r\cos u\sin v, -r\sin u\sin v, r\cos v)$$

The cross product is

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} &= \det\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -r\sin u\cos v & r\cos u\cos v & 0 \\ -r\cos u\sin v & -r\sin u\sin v & r\cos v\end{pmatrix} \\ &= (r^2\cos u\cos^2 v)\mathbf{i} + (r^2\sin u\cos^2 v)\mathbf{j} + (r^2\sin v\cos v)\mathbf{k}.\end{aligned}$$

So

$$\begin{aligned}
a(S) &= \iint_T \sqrt{r^4(\cos^2 u \cos^4 v + \sin^2 u \cos^4 v + \sin^2 v \cos^2 v)}\, \mathrm{d}u\mathrm{d}v \\
&= r^2 \int_{-\pi/2}^{\pi/2} \left(\int_0^{2\pi} \sqrt{\cos^4 v + \sin^2 v \cos^2 v}\, \mathrm{d}u \right) \mathrm{d}v \\
&= r^2 \int_{-\pi/2}^{\pi/2} \left(\int_0^{2\pi} |\cos v|\, \mathrm{d}u \right) \mathrm{d}v \\
&= 2\pi r^2 \int_{-\pi/2}^{\pi/2} |\cos v|\, \mathrm{d}v = 4\pi r^2. \qquad \square
\end{aligned}$$

Since this is measuring a geometric quantity of S, it should not depend on the way that S was parametrized, much like the arc length of a curve does not depend on its parametrization. We'll see why that is later. Using this definition, we can generalize the definition of a line integral with respect to arc length. Let $f\colon S \to \mathbf{R}$ be a bounded function on S. We define the **surface integral of f (with respect to surface area)** to be

$$\iint_{\mathbf{r}(T)} f\, \mathrm{d}S := \iint_T f(\mathbf{r}(u,v)) \left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| \mathrm{d}u\mathrm{d}v$$

whenever the right side exists (we have discussed conditions for that to be the case, for example if f and $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ are continuous, or they are discontinuous on a set of content zero).

So far, this is different from how we treated line integrals: recall that we took the line integral of a vector field, and not a scalar field. Let $N = \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ be the fundamental vector product (which is normal to S at every point) and let $\mathbf{n} = N/\|N\|$ be the unit length vector product. Given a vector field $F\colon S \to \mathbf{R}^3$, the dot product $F \cdot \mathbf{n}$ is a scalar field, so we can take its integral:

$$\iint_S F \cdot \mathbf{n}\, \mathrm{d}S = \iint_T F(\mathbf{r}(u,v)) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \mathrm{d}u\mathrm{d}v.$$

This is analogous to how we defined line integrals (thinking of the fundamental vector product as taking the role of the derivative of the parametrization). It perhaps would have been less confusing to do this in the same order as was done in line integrals, but we're following Apostol's exposition to maintain consistency. In that sense, the integral above might be denoted $\iint F \cdot \mathbf{r}$, but we won't introduce that notation. To make it easier to visualize, here is a table:

Line	Surface
$\int$(vector field) $\cdot\, \mathrm{d}\alpha$	$\int$(vector field) $\cdot\, \mathbf{n}\, \mathrm{d}S$
$\int$(scalar field) $\cdot\, \mathrm{d}s$	$\int$(scalar field) $\mathrm{d}S$

4.4. **Change of parametrization.** Suppose we are given a parametrized surface $\mathbf{r}\colon A \to \mathbf{R}^3$ where A is a region in the u,v-plane. Suppose we are given another region B in the s,t-plane and a continuously differentiable function $G(s,t) = (U(s,t), V(s,t))$ that sends A to B, i.e.,

$$A = \{(U(s,t), V(s,t)) \mid (s,t) \in B\}.$$

Furthermore, we assume that G is injective, i.e., $G(s,t) = G(s_0, t_0)$ implies that $(s,t) = (s_0, t_0)$. If we define $\rho(s,t) = \mathbf{r}(G(s,t))$, then $\rho(B)$ is another parametrization for the surface $\mathbf{r}(A)$. The parametrizations $\mathbf{r}$ and ρ are called **smoothly equivalent**.

Lemma 4.8. *For any $(a,b) \in B$, we have*

$$\frac{\partial \rho}{\partial s}(a,b) \times \frac{\partial \rho}{\partial t}(a,b) = \frac{\partial(U,V)}{\partial(s,t)} \left(\frac{\partial \mathbf{r}}{\partial u}(U(a,b), V(a,b)) \times \frac{\partial \mathbf{r}}{\partial v}(U(a,b), V(a,b)) \right).$$

Proof. By the chain rule, we have

$$\frac{\partial \rho}{\partial s} = \frac{\partial \mathbf{r}}{\partial u}\frac{\partial U}{\partial s} + \frac{\partial \mathbf{r}}{\partial v}\frac{\partial V}{\partial s}, \qquad \frac{\partial \rho}{\partial t} = \frac{\partial \mathbf{r}}{\partial u}\frac{\partial U}{\partial t} + \frac{\partial \mathbf{r}}{\partial v}\frac{\partial V}{\partial t}$$

where $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ are evaluated at $(U(a,b), V(a,b))$ and all the others are evaluated at (a,b). The rest follows if we remember that $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$ and $\mathbf{x} \times \mathbf{x} = 0$ for any vectors $\mathbf{x}, \mathbf{y}$ (and $\frac{\partial U}{\partial s}$ and $\frac{\partial U}{\partial t}$ evaluated at (a,b) are just numbers):

$$\frac{\partial \rho}{\partial s} \times \frac{\partial \rho}{\partial t} = \left(\frac{\partial U}{\partial s}\frac{\partial U}{\partial t} - \frac{\partial V}{\partial s}\frac{\partial V}{\partial t} \right) \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = \frac{\partial(U,V)}{\partial(s,t)} \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}. \qquad \square$$

Theorem 4.9. *Notation as above. Let $f\colon S \to \mathbf{R}$ be a bounded function such that $\iint_{\mathbf{r}(A)} f\,\mathrm{d}S$ exists. Assume also that $\frac{\partial(U,V)}{\partial(s,t)}$ is always positive. Then $\iint_{\rho(B)} f\,\mathrm{d}S$ also exists and the two integrals are equal.*

Proof. We have

$$\begin{aligned}
\iint_{\mathbf{r}(A)} f\,\mathrm{d}S &= \iint_A f(\mathbf{r}(u,v)) \left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| \mathrm{d}u\mathrm{d}v && \text{(by definition)} \\
&= \iint_B f(\mathbf{r}(G(s,t))) \left\| \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) (U(s,t), V(s,t)) \right\| \frac{\partial(U,V)}{\partial(s,t)} \mathrm{d}s\mathrm{d}t && \text{(change of variables)} \\
&= \iint_B f(\rho(s,t)) \left\| \frac{\partial \rho}{\partial s} \times \frac{\partial \rho}{\partial t} \right\| \mathrm{d}s\mathrm{d}t && \text{(Lemma 4.8)} \\
&= \iint_{\rho(B)} f\,\mathrm{d}S. && \square
\end{aligned}$$

4.5. **Curl and divergence.** Let $F\colon S \to \mathbf{R}^3$ be a differentiable vector field defined on some region S. Write the components of F as (P, Q, R), i.e., $F = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$. The **curl** of F, denoted $\operatorname{curl} F$, is another vector field defined as follows:

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

Here is a heuristic for remembering it. If we think of $\frac{\partial P}{\partial x}$, etc. as the "product" of $\frac{\partial}{\partial x}$ and P, then $\operatorname{curl} F$ can be written as a determinant:

$$\operatorname{curl} F = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{pmatrix}.$$

Going with this a little further, we can think of the gradient ∇ as the vector $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$, and then the above is simply a cross product:

$$\operatorname{curl} F = \nabla \times F.$$

We have to keep in mind that this has a very specific meaning that we have just introduced and this is not the same as we've been using cross products. However, this shorthand is very helpful for remembering the formula so the abuse of notation is worth it.

The **divergence** of F is a scalar field, denoted $\operatorname{div} F$, and is defined by

$$\operatorname{div} F = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

Following our new notation, we can write this as a dot product:

$$\operatorname{div} F = \nabla \cdot F.$$

While we're at it, if $\varphi \colon S \to \mathbf{R}$ is a scalar field, then we can interpret the gradient as a product $\nabla\varphi$, which is already the notation we use.

Remark 4.10. We can interpret Theorem 2.13 as saying that if S is an open convex set (or the generalization we discussed there), then $F \colon S \to \mathbf{R}^3$ is a gradient if and only if $\operatorname{curl} F = 0$. $\square$

Here are some basic properties of div and curl which you'll prove in homework.

Theorem 4.11. *Let $S \subset \mathbf{R}^3$, let a, b be real numbers, let $F, G \colon S \to \mathbf{R}^3$ be vector fields with continuous second partial derivatives, and let $\varphi \colon S \to \mathbf{R}$ be a scalar field. The following properties hold:*

$$\begin{aligned}
\operatorname{div}(aF + bG) &= a \operatorname{div} F + b \operatorname{div} G \\
\operatorname{curl}(aF + bG) &= a \operatorname{curl} F + b \operatorname{curl} G \\
\operatorname{div}(\operatorname{curl} F) &= 0 \\
\operatorname{div}(\varphi F) &= \varphi \operatorname{div} F + \nabla\varphi \cdot F \\
\operatorname{curl}(\varphi F) &= \varphi \operatorname{curl} F + \nabla\varphi \times F.
\end{aligned}$$

The last two can be rewritten with our special notation:

$$\begin{aligned}
\nabla \cdot (\varphi F) &= \varphi(\nabla \cdot F) + \nabla\varphi \cdot F \\
\nabla \times (\varphi F) &= \varphi(\nabla \times F) + (\nabla\varphi) \times F.
\end{aligned}$$

This looks formally like the product rule for derivatives, and gives a way to remember these formulas more easily.

4.6. **Stokes' theorem.**

Theorem 4.12 (Stokes' theorem)**.** *Let $\mathbf{r} \colon T \to S$ be a smooth, simple parametric surface, and assume that the boundary of T is a piecewise smooth Jordan curve Γ and assume that $\mathbf{r}$ has continuous second-order partial derivatives in some neighborhood of $T \cup \Gamma$. Let $C = \mathbf{r}(\Gamma)$ be the boundary of S, and let $F \colon S \to \mathbf{R}^3$ be a continuously differentiable vector field. Then*

$$\iint_S (\operatorname{curl} F) \cdot \mathbf{n} \, \mathrm{d}S = \int F \cdot \mathrm{d}\alpha,$$

where β is a counterclockwise simple parametrization of Γ, and $\alpha = \mathbf{r} \circ \beta$ is a simple parametrization of C.

For shorthand, we might write this as $\iint_S \operatorname{curl} F = \int_{\partial S} F$, but we have to remember all of the conventions in place about parametrizations. If we think of $\operatorname{curl} F$ as a sort of derivative for F, then we can view this as a version of the second fundamental theorem of calculus. In fact, Green's theorem gives such a version for surfaces in $\mathbf{R}^2$, and this is just an extension to surfaces in $\mathbf{R}^3$.

Proof. The idea is to replace both integrals with integrals over T and Γ, respectively, so that the desired equality follows from Green's theorem.

First assume that $F(x, y, z) = (P(x, y, z), 0, 0)$ for a scalar field P, i.e., assume that the second and third components of F are 0. In that case, $\operatorname{curl} F = (0, \frac{\partial P}{\partial z}, -\frac{\partial P}{\partial y})$. Define a function $\varphi\colon T \to \mathbf{R}$ by

$$\varphi(u, v) = P(X(u, v), Y(u, v), Z(u, v)).$$

Then we have the identity (deferred to homework):

$$\left(0, \frac{\partial P}{\partial z}, -\frac{\partial P}{\partial y}\right) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) = \frac{\partial}{\partial u}\left(\varphi \frac{\partial X}{\partial v}\right) - \frac{\partial}{\partial v}\left(\varphi \frac{\partial X}{\partial u}\right).$$

So the left hand side of the desired identity is

$$\iint_T \left(0, \frac{\partial P}{\partial z}, -\frac{\partial P}{\partial y}\right) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) \mathrm{d}u\mathrm{d}v = \iint_T \frac{\partial}{\partial u}\left(\varphi \frac{\partial X}{\partial v}\right) - \frac{\partial}{\partial v}\left(\varphi \frac{\partial X}{\partial u}\right) \mathrm{d}u\mathrm{d}v.$$

Using Green's theorem, this becomes

$$\int \left(\varphi \frac{\partial X}{\partial u}, \varphi \frac{\partial X}{\partial v}\right) \cdot \mathrm{d}\beta.$$

Finally, we have $\alpha(t) = (X(\beta(t)), Y(\beta(t)), Z(\beta(t))$, so the right hand side of the desired identity is (let $[a, b]$ be the domain of the parametrization of β)

$$\int F \cdot \mathrm{d}\alpha = \int_a^b P(\mathbf{r}(\beta(t))) \frac{\mathrm{d}}{\mathrm{d}t}(X(\beta(t)))\, \mathrm{d}t.$$

But this last integral is the same as the previous one, so we conclude that

$$\iint_S (\operatorname{curl} F) \cdot \mathbf{n}\, \mathrm{d}S = \int F \cdot \mathrm{d}\alpha$$

in this special case. By a similar calculation, we can show the same identity when $F = (0, Q, 0)$ and when $F = (0, 0, R)$. If you add together the identities from these 3 special cases, you get the identity for the general case. $\square$

We can use this theorem in two ways: either to turn a line integral into a surface integral, or the other way around.

Example 4.13. Let C be the curve of intersection of the the plane $y + z = 2$ and the cylinder $x^2 + y^2 = 1$ with the orientation counterclockwise when viewed from above. If we want to evaluate a line integral $\int f \cdot \mathrm{d}\alpha$ where α parametrizes C counterclockwise, then we can instead calculate $\int_S (\operatorname{curl} F) \cdot \mathbf{n}\ \mathrm{d}S$ where S is any surface whose boundary is C (but make sure to keep track of the orientation). There are infinitely many choices for such a surface, one such obvious choice is the ellipse E that it fills out which is described by

$$E = \{(x, y, z) \mid y + z = 2,\ x^2 + y^2 \leq 1\}.$$

This is the graph of a function, so has an easy parametrization. We take $T = \{(u,v) \mid u^2+v^2 \le 1\}$ and $\mathbf{r}(u,v) = (u,v,2-v)$. The boundary of T is the unit circle in the u,v-plane, and if we parametrize it counterclockwise and then apply $\mathbf{r}$, we see that we're going counterclockwise when looking from above, so we have the right orientation (if we didn't, we could just multiply by -1 at the end of our calculation).

So we get

$$\int f \cdot \mathrm{d}\alpha = \iint_E (\operatorname{curl} f) \cdot \mathbf{n}\, \mathrm{d}S = \iint_T (\operatorname{curl} f)(\mathbf{r}(u,v)) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \mathrm{d}u\mathrm{d}v.$$

We simplify this: $\frac{\partial \mathbf{r}}{\partial u} = (1,0,0)$, $\frac{\partial \mathbf{r}}{\partial v} = (0,1,-1)$, so the cross product is $(0,1,1)$.

For concreteness, let's pick $f(x,y,z) = (-y^2, x, z^2)$. Then $\operatorname{curl} f = (0,0,1+2v)$, so we can continue to simplify:

$$\iint_T (0,0,1+2v) \cdot (0,1,1)\, \mathrm{d}u\mathrm{d}v = \iint_T (1+2v)\, \mathrm{d}u\mathrm{d}v.$$

Best to use polar coordinates now:

$$\int_0^{2\pi} \int_0^1 (1 + 2r\sin\theta) r\, \mathrm{d}r\mathrm{d}\theta = \cdots = \pi. \qquad \square$$

4.7. **Uncurling a vector field.** In the previous example, we use Stokes' theorem to convert a line integral into a surface integral. This is backwards from how we use the fundamental theorem of calculus, i.e., moving down in the dimension should be simpler rather than harder. To convert surface integrals into line integrals, one difficulty we face is that we need to solve the equation $\operatorname{curl} F = G$ given G, analogous to how we solve the equation $f' = g$ when we solve 1-dimensional integrals.

As explained before, if $\operatorname{curl} F$ has continuous partial derivatives, then we have $\operatorname{div}(\operatorname{curl} F) = 0$, so given G, we can't solve this equation unless $\operatorname{div} G = 0$. In some cases, this is actually enough to guarantee that F exists. You can think of this as analogous to the statement that gradient satisfies certain equalities among its partial derivatives, and that in some cases (like convex regions), these equalities also guarantee that a vector field is actually a gradient. Here's one such statement (recall that an interval in $\mathbf{R}^3$ is a rectangular prism, i.e., a region of the form $[a_1,b_1] \times [a_2,b_2] \times [a_3,b_3]$):

Theorem 4.14. *Let G be a continuously differentiable vector field on an interval in $\mathbf{R}^3$. Then there exists a vector field F with $\operatorname{curl} F = G$ if and only if $\operatorname{div} G = 0$. Furthermore, we can define F as follows. Fix a point (x_0,y_0,z_0) in the interval. Then we can take $F = (0,M,N)$ where*

$$M(x,y,z) = \int_{x_0}^{x} G_3(t,y,z)\, \mathrm{d}t - \int_{z_0}^{z} G_1(x_0,y,u)\, \mathrm{d}u$$
$$N(x,y,z) = -\int_{x_0}^{x} G_2(t,y,z)\, \mathrm{d}t.$$

If our vector field is defined on all of $\mathbf{R}^3$, the above still works since we can pick a point (x_0,y_0,z_0) once and for all and verify that the formulas work in any interval containing this point (and hence everywhere). For simplicity, we can take the point $(0,0,0)$.

Proof. We already know that the existence of F implies that $\operatorname{div} G = 0$, so we just need to verify that the formula above works under the assumption that $\operatorname{div} G = 0$. We will use

Theorem 2.14, which says that we can differentiate the integrals in the above formula by differentiating the integrand, and we will also use both fundamental theorems of calculus for 1-dimensional integrals. With our formula, we have $\operatorname{curl} F = (\frac{\partial N}{\partial y} - \frac{\partial M}{\partial z}, -\frac{\partial N}{\partial x}, \frac{\partial M}{\partial x})$. Now we verify:

$$\begin{aligned}\frac{\partial N}{\partial y} - \frac{\partial M}{\partial z} &= -\int_{x_0}^{x} \frac{\partial G_2}{\partial y}(t,y,z)\,\mathrm{d}t - \int_{x_0}^{x} \frac{\partial G_3}{\partial z}(t,y,z)\,\mathrm{d}t + G_1(x_0,y,z)\\ &= \int_{x_0}^{x} \frac{\partial G_1}{\partial x}(t,y,z)\,\mathrm{d}t + G_1(x_0,y,z)\\ &= G_1(x,y,z) - G_1(x_0,y,z) + G_1(x_0,y,z) = G_1(x,y,z)\end{aligned}$$

where in the second line we used that $\operatorname{div} G = 0$. We also have

$$-\frac{\partial N}{\partial x} = G_2(x,y,z), \qquad \frac{\partial M}{\partial x} = G_3(x,y,z)$$

by the first fundamental theorem of calculus. So we conclude that $\operatorname{curl} F = G$. □

Example 4.15. Let $G(x,y,z) = (x,y,-2z)$, so $\operatorname{div} G = 0$. The above theorem tells us that $G = \operatorname{curl} F$ where $F = (0, M, N)$ and

$$\begin{aligned}M(x,y,z) &= \int_0^x -2z\,\mathrm{d}t - \int_0^z 0\,\mathrm{d}u = -2xz,\\ N(x,y,z) &= -\int_0^x y\,\mathrm{d}t = -xy.\end{aligned}$$

You can verify directly that $\operatorname{curl}(0,-2xz,-xy) = (x,y,-2z)$. □

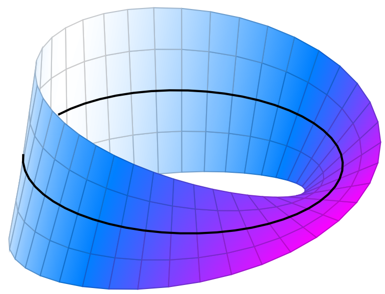
How unique is the solution? If $\operatorname{curl} F = \operatorname{curl} H = G$, then $\operatorname{curl}(F - H) = 0$, and since intervals (and $\mathbf{R}^3$) are convex, by Remark 4.10, there exists a scalar field φ such that $\nabla\varphi = F - H$. Hence any two solutions to $\operatorname{curl} F = G$ differ by the gradient of a scalar field. So the most general solution to $\operatorname{curl} F = G$ is

$$F = \left(\frac{\partial\varphi}{\partial x}, M + \frac{\partial\varphi}{\partial y}, N + \frac{\partial\varphi}{\partial z}\right)$$

where M, N are defined as in the previous theorem.

4.8. **The divergence theorem.** We will need the notion of an orientable surface in $\mathbf{R}^3$. Given a parametrization, we have discussed the unit normal vector of the fundamental vector product. There are actually 2 unit normal vectors: the one we discussed as well as its negative. We'll say that a surface is **orientable** if it is possible to choose a unit normal vector at all points in a continuous way. If it is not possible, we call it **non-orientable**. For example, the 2-sphere is orientable: we can choose the unit normals to be pointing "outside". An example of a non-orientable surface is the Möbius strip, which is drawn as follows and has the following parametrization ($-1/2 \le r \le 1/2$ and $0 \le \theta \le 2\pi$):

$$X(r,\theta) = 2\cos\theta + r\cos(\theta/2), \qquad Y(r,\theta) = 2\sin\theta + r\cos(\theta/2), \qquad Z(r,\theta) = r\sin(\theta/2)$$



If you've never seen this before, take a sheet of paper, wrap the ends around and rotate one of the sides and tape them together.

Finally, a surface is **closed** if its complement is open, it is bounded, and it has no boundary.

Theorem 4.16 (Divergence theorem)**.** *Let V be a bounded 3-dimensional region in $\mathbf{R}^3$ whose boundary is an orientable closed surface S, and let $\mathbf{n}$ be the unit normal vector field on S which points away from V. If $F\colon V \to \mathbf{R}^3$ is a continuously differentiable vector field, then*

$$\iint_V (\operatorname{div} F)\, \mathrm{d}x\mathrm{d}y\mathrm{d}z = \iint_S F \cdot \mathbf{n}\, \mathrm{d}S.$$

Proof. The proof of the divergence theorem can be done by a strategy similar to the proof we gave for Green's theorem. First one assumes that the region V is simultaneously of type I, type II, and what you would call type III (for 3-dimensions). The point is to first consider vector fields F where only one of its component functions is nonzero and then to explicitly compare the two sides of the equations. This gives 3 different special identities which can be added together to get the general identity. More general regions can be handled by cutting them into pieces which are simultaneously of types I, II, and III. We won't go through the details, though you can find them in §12.19 of Apostol. ☐

Again, we can think of the divergence of a vector field as a sort of derivative, so that this can be thought of as a 3-dimensional version of the second fundamental theorem of calculus.

Remark 4.17. The divergence theorem can be used to prove an alternative formula for the divergence of a vector field which can then be used to give a physical interpretation. See §12.20 of Apostol for details, though we won't discuss it in this class. ☐

Example 4.18. Consider the region V bounded by the graph of $z = 1 - x^2$ and the planes $z = 0$, $y = 0$, and $y + z = 2$ and let S be the boundary. There are 4 pieces of S, so evaluating surface integrals directly requires a bunch of work. Alternatively, we can use the divergence theorem to get

$$\iint_S F \cdot \mathbf{n}\, \mathrm{d}S = \iint_V \operatorname{div} F\, \mathrm{d}x\mathrm{d}y\mathrm{d}z$$

where $\mathbf{n}$ is the outwards pointing normal. The region V can be described by 3 inequalities:

$$V = \{(x, y, z) \mid -1 \le x \le 1,\ 0 \le z \le 1 - x^2,\ 0 \le y \le 2 - z\},$$

so we can setup the triple integral as follows:

$$\int_{-1}^{1} \left(\int_0^{1-x^2} \left(\int_0^{2-z} \operatorname{div} F\, \mathrm{d}y \right) \mathrm{d}z \right) \mathrm{d}x. \qquad \square$$

Since integrating the function 1 on a 3-dimensional region gives us its volume, we can use the divergence theorem to turn this problem into a surface integral since $\mathrm{div}(ax, by, cz) = 1$ whenever $a + b + c = 1$. So if V has boundary S, then

$$\mathrm{vol}(V) = \iiint_V \mathrm{d}x\mathrm{d}y\mathrm{d}z = \iint_S (ax, by, (1-a-b)z) \cdot \mathbf{n}\, \mathrm{d}S$$

for any choice of $a, b \in \mathbf{R}$ and where $\mathbf{n}$ is the outwards pointing normal. I don't know of a good example where the surface integral is actually less work, but let's see how this would work for the ball of radius R.

Example 4.19. If V is the ball of radius R, then its boundary S is the sphere of radius R. We can parametrize S by taking $T = \{(u, v) \mid 0 \le u \le 2\pi,\ -\pi/2 \le v \le \pi/2\}$ and

$$\mathbf{r}(u, v) = (R \cos u \cos v, R \sin u \cos v, R \sin v).$$

In Example 4.7, we computed the fundamental vector product to be

$$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = (R^2 \cos u \cos^2 v, R^2 \sin u \cos^2 v, R^2 \sin v \cos v).$$

Let's take $F = \frac{1}{3}(x, y, z)$. Then

$$\begin{aligned}
\mathrm{vol}(V) &= \frac{1}{3} \iint_S (x, y, z) \cdot \mathbf{n}\, \mathrm{d}S \\
&= \iint_T \frac{R^3}{3} (\cos^2 u \cos^3 v + \sin^2 u \cos^3 v + \sin^2 v \cos v)\, \mathrm{d}u\mathrm{d}v \\
&= \frac{R^3}{3} \iint_T \cos v\, \mathrm{d}u\mathrm{d}v \\
&= \frac{R^3}{3} \int_{-\pi/2}^{\pi/2} \left(\int_0^{2\pi} \cos v\, \mathrm{d}u \right) \mathrm{d}v \\
&= \frac{2\pi R^3}{3} \int_{-\pi/2}^{\pi/2} \cos v\, \mathrm{d}v = \frac{4\pi R^3}{3}.
\end{aligned}$$

□

5. Linear differential equations

5.1. **Definitions.** In our discussion of differential equations, an **interval** shall mean something of the form (a, b) where a, b are either real numbers of $\pm\infty$. We will denote it by J throughout. If both a, b are numbers, then J is **bounded**, otherwise, it is **unbounded**.

Let $\mathscr{C}^n(J)$ denote the set of functions $f \colon J \to \mathbf{R}$ such that the derivatives $f', f'', \dots, f^{(n-1)}, f^{(n)}$ all exist and are continuous on J. Some special cases: if $n = 0$, then $\mathscr{C}^0(J)$ is the set of continuous functions on J (with no requirement on the existence of derivatives) and if $n = \infty$, then $\mathscr{C}^\infty(J)$ is the set of functions with all derivatives. Note that $\mathscr{C}^n(J)$ is a vector space with the usual addition of functions and scalar multiplication by real numbers. We will use the convention that $f^{(0)} = f$, i.e., the zeroth derivative of a function is the function itself.

Let $P_1(x), \dots, P_n(x), R(x) \in \mathscr{C}^0(J)$. We consider the differential equation

$$y^{(n)} + P_1(x) y^{(n-1)} + \cdots + P_n(x) y = R(x).$$

The n here is the **order** of the equation, and we say first-order, second-order, etc. Solving the differential equation means finding a function f such that substituting $y = f(x)$ makes the above equality true. Associated to the equation above we have a linear transformation

$$L\colon \mathscr{C}^n(J) \to \mathscr{C}^0(J)$$
$$L(f) = f^{(n)} + P_1 f^{(n-1)} + \cdots + P_n f$$

since $L(f+g) = L(f) + L(g)$ and $L(cf) = cL(f)$ for any real number c. Alternatively, we are trying to find some f such that $L(f) = R$. The derivative operator will be denoted by D, i.e., $D(f) = f'$, so we can write $L = D^n + P_1 D^{n-1} + \cdots + P_n$.

The equation $L(y) = 0$ is called the **homogeneous equation** of the differential equation above, and $L(y) = R$ is the **nonhomogeneous equation** when R is not the zero function.

The solutions to $L(y) = 0$ is the nullspace of L (by definition). We will use the word **kernel** in place of nullspace, and denote it by $\ker L$. As you have seen, $\ker L$ is a subspace of $\mathscr{C}^n(J)$.

5.2. **Existence-uniqueness of solutions.** Throughout this section, we keep the notation from the previous section.

Theorem 5.1. *Pick $x_0 \in J$ and let $k_0, \dots, k_{n-1}$ be real numbers. Then there is exactly one function $y = f(x)$ such that $L(y) = 0$ and $f^{(i)}(x_0) = k_i$ for $i = 0, \dots, n-1$.*

The conditions $f^{(i)}(x_0) = k_i$ are called the **initial conditions**. The proof is discussed in Chapter 7 of Apostol, so we will postpone it until we discuss that material. Using it, we can prove the dimensionality theorem:

Theorem 5.2 (Dimensionality theorem)**.** $\dim \ker L = n$.

Proof. Pick $x_0 \in J$. The previous theorem tells us that each $f \in \ker L$ is determined by the n numbers $f(x_0), f'(x_0), \dots, f^{(n-1)}(x_0)$. In particular, the linear map $\ker L \to \mathbf{R}^n$ which sends f to this n-tuple is an isomorphism of vector spaces. □

Corollary 5.3. *If $u_1, \dots, u_n \in \mathscr{C}^n(J)$ are linearly independent elements of $\ker L$, then every $f \in \ker L$ has a unique expression*

$$f = c_1 u_1 + \cdots + c_n u_n$$

for some real numbers $c_1, \dots, c_n$.

Proof. Since $\dim \ker L = n$, any collection of n linearly independent elements is a basis. □

Now consider the original problem of finding solutions for $L(f) = R$.

Corollary 5.4. *Let $u_1, \dots, u_n$ be a basis for $\ker L$ and let y_1 satisfy $L(y_1) = R$. Then all solutions to $L(f) = R$ can be uniquely written in the form*

$$y_1 + c_1 u_1 + \cdots + c_n u_n$$

for real numbers $c_1, \dots, c_n$, and all such expressions are solutions.

Proof. Let y satisfy $L(y) = R$. Then $L(y - y_1) = 0$, so $y - y_1 \in \ker L$ and hence $y - y_1$ can be written uniquely as a linear combination of the $u_1, \dots, u_n$. Conversely, $L(y_1 + c_1 u_1 + \cdots + c_n u_n) = L(y_1) = R$ since $c_1 u_1 + \cdots + c_n u_n \in \ker L$. □

Hence to solve the differential equation $L(f) = R$, there are two parts. First, we find a basis for $\ker L$. Second, we find a single solution to $L(f) = R$. In this case, we call y_1 a particular solution, and $y_1 + c_1 u_1 + \cdots + c_n u_n$ the general solution. We now discuss how to do the first step in a special case.

5.3. **The constant-coefficient case.** The simplest case to consider is when all of the functions P_i are constant. To emphasize that they are constants, we will write a_i instead of P_i. In that case, we call L a **constant-coefficient operator**.

Given a constant-coefficient operator $A = D^n + a_1 D^{n-1} + \cdots + a_n$, we define its **characteristic polynomial** $p_A(t) = t^n + a_1 t^{n-1} + \cdots + a_n$. We now work with infinitely differentiable functions, i.e., the space $\mathscr{C}^\infty(J)$. Note that a constant-coefficient operator A defines a linear map $\mathscr{C}^\infty(J) \to \mathscr{C}^\infty(J)$. In that case, it makes sense to compose two constant-coefficient operators A, B, and we have $AB = BA$, i.e., they commute.

Theorem 5.5. *Let A, B be constant-coefficient operators. Then*

(1) $p_A(t) = p_B(t)$ if and only if $A = B$,
(2) $p_{A+B} = p_A + p_B$,
(3) $p_{AB} = p_A p_B$,
(4) $p_{cA} = cp_A$ for any real number c.

Proof. Certainly $p_A = p_B$ implies that $A = B$. Conversely, suppose that $A = B$, i.e., $A(f) = B(f)$ for all $f \in \mathscr{C}^\infty(J)$. In particular, consider $f(x) = e^{rx}$ for some real number r. Then $A(f) = p_A(r)e^{rx}$ and $B(f) = p_B(r)e^{rx}$ which means that $p_A(r) = p_A(r)$ for all real numbers r, so $p_A(t)$ and $p_B(t)$ are the same polynomial.

The other properties follow from the definitions, and the proof is omitted. □

The polynomial $p_A(t)$ can be factored as a product of linear factors over the complex numbers (by the fundamental theorem of algebra). Note that if λ is a root of $p_A(t)$, then so is its complex conjugate $\overline{\lambda}$ (since the coefficients of $p_A(t)$ are real), so the non-real roots of $p_A(t)$ come in complex conjugate pairs. Note that $(t-\lambda)(t-\overline{\lambda})$ has real coefficients. So we conclude that any $p_A(t)$ can always be factored as a product of linear polynomials and quadratic polynomials all with real coefficients. Furthermore, we may assume that all of the quadratic polynomials have no real solutions, which is equivalent to its discriminant being negative (recall the discriminant of $at^2 + bt + c$ is $b^2 - 4ac$).

Every polynomial is the characteristic polynomial of some constant-coefficient operator, so we can write the factorization as

$$p_A(t) = p_{A_1}(t) \cdots p_{A_k}(t)$$

where each A_i is either a first-order or second-order differential equation. Note that this implies an identity

$$A = A_1 \cdots A_k.$$

Theorem 5.6. *Suppose we have a factorization of a constant-coefficient operator as a product of other constant-coefficient operators*

$$A = A_1 \cdots A_k.$$

Then $\ker A_i \subseteq \ker A$ *for all* $i = 1, \dots, k$.

Proof. Since constant-coefficient operators commute, for any i we can write $A = BA_i$ where B is the product of the A_j where $j \neq i$. If $f \in \ker A_i$, then we have $A(f) = B(A_i(f)) = B(0) = 0$, so $f \in \ker A$. □

We now solve the homogeneous equation of the general constant-coefficient operator. We first factor

$$p_A(t) = p_{A_1}(t)^{m_1} \cdots p_{A_k}(t)^{m_k}$$

where the p_{A_i} are all distinct polynomials with leading coefficient 1, which are either linear, or quadratic with no real roots. The m_i are positive integers, called the **multiplicities** of the p_{A_i}. We first solve the cases $p_{A_i}(t)^{m_i}$ and then combine the answers.

Lemma 5.7. *Given a real number r, the functions $u_1(x) = e^{rx}, u_2(x) = xe^{rx}, \cdots, u_m(x) = x^{m-1}e^{rx}$ form a basis for* $\ker((D-r)^m)$.

Proof. First, we show that $(D-r)^m u_i(x) = 0$ for $i = 1, \dots, m$. We have

$$(D-r)x^i e^{rx} = (x^{i-1}e^{rx} + rx^i e^{rx}) - rx^i e^{rx} = x^{i-1}e^{rx}.$$

From this, we see that $(D-r)^m(x^i e^{rx}) = 0$ if $0 \le i \le m-1$.

Now we claim that $u_1, \dots, u_m$ are linearly independent. Suppose that there are real numbers $c_1, \dots, c_m$ such that

$$c_1 u_1 + \cdots c_m u_m = 0.$$

Divide by e^{rx} to get $c_1 + c_2 x + \cdots + c_m x^{m-1} = 0$. This is true for all x, which implies that $c_1 + c_2 x + \cdots + c_m x^{m-1}$ is the zero polynomial, i.e., all $c_i = 0$, which proves linear independence.

Finally, the dimensionality theorem implies that $\dim\ker(D-r)^m = m$, so these linearly independent elements must form a basis. □

Lemma 5.8. *Suppose that $b^2 - 4c < 0$ and let $\alpha + i\beta$ and $\alpha - i\beta$ be the complex solutions to $t^2 + bt + c = 0$. For $j = 1, \dots, m$, define*

$$u_j(x) = x^{j-1}e^{\alpha x}\cos(\beta x), \qquad v_j(x) = x^{j-1}e^{\alpha x}\sin(\beta x).$$

Then $u_1, \dots, u_m, v_1, \dots, v_m$ is a basis for $\ker(D^2 + bD + c)^m$.

Proof. While this can be done without using complex numbers, it is easier if we take advantage of a few identities:

$$e^{ix} = \cos(x) + i\sin(x), \qquad \cos(x) = \frac{1}{2}(e^{ix} + e^{-ix}), \qquad \sin(x) = \frac{1}{2i}(e^{ix} - e^{-ix}).$$

Adapting the previous result, a basis for $\ker(D - (\alpha + i\beta))^m$ is given by

$$e^{(\alpha+i\beta)x}, xe^{(\alpha+i\beta)x}, \dots, x^{m-1}e^{(\alpha+i\beta)x}.$$

Similarly, for $\ker(D - (\alpha - i\beta))^m$, a basis is given by

$$e^{(\alpha-i\beta)x}, xe^{(\alpha-i\beta)x}, \dots, x^{m-1}e^{(\alpha-i\beta)x}.$$

Next, we have

$$e^{(\alpha+i\beta)x} = e^{\alpha x}(\cos(\beta x) + i\sin(\beta x)),$$

so

$$x^j e^{\alpha x}\cos(\beta x) = \frac{1}{2}(x^j e^{(\alpha+i\beta)x} + x^j e^{(\alpha-i\beta)x}),$$
$$x^j e^{\alpha x}\sin(\beta x) = \frac{1}{2i}(x^j e^{(\alpha+i\beta)x} - x^j e^{(\alpha-i\beta)x}).$$

So $u_j, v_j \in \ker(D^2 + bD + c)^m$. We'll omit the check that these are linearly independent (I'll outline one way in homework). Given that fact, they form a basis by the dimensionality theorem. □

For the general case, we take the bases for each $\ker(A_i^{m_i})$ and combine them to get a set of solutions for $\ker(A)$. In fact, these solutions are all linearly independent. We won't prove this for the sake of avoiding heavy amounts of notation, but one way to do it would be to argue that if a linear combination with coefficients c_i were 0, then the corresponding Taylor series of the linear combination also has to be 0. Then one uses that the Taylor series of a sum is the sum of the corresponding Taylor series, and this will force all of the c_i to be 0 (there are other ways to do it).

Example 5.9. Suppose that

$$p_A(t) = (t+4)^2(t-3)^3(t^2-2t+5)^2.$$

The roots of $t^2 - 2t + 5$ are $1 \pm 2i$, so a basis for $\ker(D^2 - 2t + 5)^2$ is given by

$$e^x \cos(2x), xe^x \cos(2x), e^x \sin(2x), xe^x \sin(2x).$$

Hence a basis for $\ker A$ is these 4 functions together with the following 5 functions:

$$e^{-4x}, xe^{-4x}, e^{3x}, xe^{3x}, x^2e^{3x}. \qquad \square$$

5.4. **Finding a particular solution.** We return to the general problem of a linear differential equation (not necessarily constant-coefficient). Recall that our equations is

$$y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_n(x)y = R(x).$$

Suppose we find a basis $u_1, \ldots, u_n$ for $\ker L$. Define the **Wronskian** to be the following $n \times n$ matrix whose (i,j) entry is the $(i-1)$st derivative of u_j:

$$W(x) = \begin{bmatrix} u_1(x) & u_2(x) & \cdots & u_n(x) \\ u_1'(x) & u_2'(x) & \cdots & u_n'(x) \\ \vdots & & & \\ u_1^{(n-1)}(x) & u_2^{(n-1)}(x) & \cdots & u_n^{(n-1)}(x) \end{bmatrix}.$$

Theorem 5.10. *$W(x)$ is invertible for all $x \in J$.*

Proof. Pick $c \in J$. Suppose that we have a solution $W(c)\alpha = 0$ for some column vector $\alpha = (\alpha_1, \ldots, \alpha_n)$. Set $f(x) = \alpha_1 u_1(x) + \cdots + \alpha_n u_n(x)$. Multiplying out $W(c)\alpha = 0$ shows that $f(c) = f'(c) = \cdots = f^{(n-1)}(c) = 0$. But $f \in \ker L$, and has the same initial conditions as the zero function, so the existence-uniqueness theorem implies f is the zero function. Since the u_i are linearly independent, we conclude that $\alpha_1 = \cdots = \alpha_n = 0$, so the only solution to $W(c)\alpha = 0$ is $\alpha = 0$. This implies that $W(c)$ is invertible. $\square$

In particular, we can define $W(x)^{-1}$ for all $x \in J$.

In what follows, the integral of a column vector is the column vector we get from taking the integral of each entry.

Theorem 5.11. *Pick $c \in J$. Let $v_1(x), \ldots, v_n(x)$ be the entries of the column vector*

$$v(x) = \int_c^x W(t)^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ R(t) \end{bmatrix} \mathrm{d}t.$$

Here, the integral of a column vector means take the integral of each entry. Then $y_1 = v_1(x)u_1(x) + \cdots + v_n(x)u_n(x)$ is a solution to the differential equation $L(f) = R$.

Proof. Take the derivative of $v(x)$ to get

$$v'(x) = W(x)^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ R(x) \end{bmatrix},$$

or equivalently,

$$W(x)v'(x) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ R(x) \end{bmatrix}.$$

It will be convenient to treat v as a row vector and $u = (u_1, \dots, u_n)$ and write $y_1 = v \cdot u$. The above matrix equation says that $v' \cdot u^{(i)} = 0$ for $i = 0, \dots, n-2$ and $v' \cdot u^{(n-1)} = R$. Now let's take derivatives of y_1.

$$\begin{aligned}
y_1' &= v' \cdot u + v \cdot u' = v \cdot u', \\
y_1'' &= v' \cdot u' + v \cdot u'' = v \cdot u'', \\
&\vdots \\
y_1^{(n-1)} &= v' \cdot u^{(n-2)} + v \cdot u^{(n-1)} = v \cdot u^{(n-1)} \\
y_1^{(n)} &= v' \cdot u^{(n-1)} + v \cdot u^{(n)} = R + v \cdot u^{(n)}
\end{aligned}$$

Hence, we have

$$y_1^{(n)} + P_1 y_1^{(n-1)} + \cdots + P_n y_1 = R + v \cdot (P_1 u^{(n)} + \cdots + P_n u) = R$$

where the last equality follows from the fact that $u_i \in \ker L$ for each i. $\square$

Example 5.12. Consider the differential equation on $J = (-\infty, \infty)$

$$y'' - 4y = e^x$$

The characteristic polynomial is $t^2 - 4 = (t-2)(t+2)$, so a basis for $\ker L$ is given by $u_1(x) = e^{2x}$, $u_2(x) = e^{-2x}$. The Wronskian is

$$W(x) = \begin{bmatrix} e^{2x} & e^{-2x} \\ 2e^{2x} & -2e^{-2x} \end{bmatrix},$$

and its inverse is

$$W(x)^{-1} = \frac{1}{4} \begin{bmatrix} 2e^{-2x} & e^{-2x} \\ 2e^{2x} & -e^{2x} \end{bmatrix}.$$

We'll take $c = 0$ and compute

$$v(x) = \frac{1}{4} \int_0^x \begin{bmatrix} 2e^{-2t} & e^{-2t} \\ 2e^{2t} & -e^{2t} \end{bmatrix} \begin{bmatrix} 0 \\ e^t \end{bmatrix} \mathrm{d}t = \frac{1}{4} \int_0^x \begin{bmatrix} e^{-t} \\ -e^{3t} \end{bmatrix} \mathrm{d}t = \frac{1}{4} \begin{bmatrix} -e^{-x} + 1 \\ (-e^{3x} + 1)/3 \end{bmatrix}.$$

So a particular solution is given by

$$v_1(x)u_1(x) + v_2(x)u_2(x) = \frac{1}{4}(1 - e^{-x})e^{2x} + \frac{1}{12}(1 - e^{3x})e^{-2x} = -\frac{e^x}{3} + \frac{e^{2x}}{4} + \frac{e^{-2x}}{12},$$

and the general solution is

$$-\frac{e^x}{3} + \frac{e^{2x}}{4} + \frac{e^{-2x}}{12} + c_1 e^{2x} + c_2 e^{-2x}$$

where c_1, c_2 are real numbers. □

Example 5.13. Consider the differential equation

$$y' + P(x)y = R(x)$$

on some integral J. The homogeneous equation can be rewritten as

$$\frac{y'}{y} = -P(x).$$

The left hand side is the logarithmic derivative of y: it is the derivative of $\ln(y)$ wherever y takes positive values. Pick a point $c \in J$. This allows us to write $\ln(y) = -\int_c^x P(t)\,\mathrm{d}t$. Set $A(x) = \int_c^x P(t)\,\mathrm{d}t$. Then the solution is any scalar multiple of $e^{-A(t)}$. Call this function $u_1(x)$. Then the Wronskian is just a 1×1 matrix with entry $u_1(x)$, so we have

$$v(x) = \int_c^x \frac{R(t)}{u_1(t)}\,\mathrm{d}t = \int_c^x R(t)e^{A(t)}\,\mathrm{d}t,$$

and so the general solution is

$$e^{-A(t)} \int_c^x R(t)e^{A(t)}\,\mathrm{d}t + \alpha e^{-A(t)}$$

where $\alpha \in \mathbf{R}$. □

6. Systems of differential equations

6.1. **Notation.** We are now concerned with systems of differential equations with several unknown functions $y_1, \ldots, y_n$. We will study the case of first-order equations, and in particular, systems which consist of n equations of the form

$$\begin{aligned} y_1' &= p_{1,1}(t)y_1 + \cdots + p_{1,n}(t)y_n + q_1(t) \\ &\vdots \\ y_n' &= p_{n,1}(t)y_1 + \cdots + p_{n,n}(t)y_n + q_n(t) \end{aligned}$$

where the $p_{i,j}(t), q(t)$ are functions on some given interval J. It will be convenient to write these in matrix notation. So we set

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \qquad P(t) = \big[p_{i,j}(t)\big]_{i,j=1}^n, \qquad Q(t) = \begin{bmatrix} q_1(t) \\ \vdots \\ q_n(t) \end{bmatrix},$$

and write our system of equations as

$$Y' = P(t)Y + Q(t).$$

Remark 6.1. A single nth order differential equation as we considered in the previous section can be encoded as a system of first order differential equations as follows. Suppose we're given the equation

$$y^{(n)} + P_1(t)y^{(n-1)} + \cdots + P_n(t)y = R(t).$$

Consider the $n-1$ equations $y_i' = y_{i+1}$ for $i = 1, \dots, n-1$ together with equation

$$y_n' = -P_1(t)y_n - P_2(t)y_{n-1} - \cdots - P_n(t)y_1 + R(t)$$

For any solution to this system, the function y_1 gives a solution to our original equation. □

So we will extend the familiar operations of calculus to matrices of functions. Integrals and derivatives are taken entrywise, meaning that the integral of a matrix is just the matrix obtained by taking the integral term by term, and similarly for derivatives. We still have the familiar sum and product rules:

$$(P+Q)' = P' + Q', \qquad (PQ)' = P'Q + PQ'.$$

6.2. **Matrix exponentials.** An important idea is the exponential of an $n \times n$ matrix, which cannot be defined entrywise (for example, we want $e^0 = I_n$ where 0 is the 0 matrix and I is $n \times n$ the identity matrix).

The useful way to define e^A, where A an $n \times n$ matrix, is to use the Taylor series expansion of e^x:

$$e^A = I_n + \sum_{k \geq 1} \frac{A^k}{k!}.$$

Since this is an infinite sum, we have to first make sense of what it means to converge. However, we can already see that it satisfies $e^0 = I_n$.

Definition 6.2. Let $(C_k)_{k \geq 1}$ be a sequence of matrices and let $c_{ij}^{(k)}$ be the (i,j) entry of C_k. We say that the infinite sum $\sum_{k \geq 1} C_k$ is **convergent** if each of the sums $\sum_{k \geq 1} c_{ij}^{(k)}$ is convergent for all pairs i, j. In that case, the (i,j) entry of $\sum_{k \geq 1} C_k$ is defined to be $\sum_{k \geq 1} c_{ij}^{(k)}$. □

Our goal now is to prove that the sum $\sum_{k \geq 1} \frac{A^k}{k!}$ is convergent for any square matrix A. To do that, we introduce the norm of a matrix.

Definition 6.3. Let A be an $m \times n$ (real or complex) matrix. The norm of A, denoted $\|A\|$, is defined to be

$$\|A\| = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{i,j}|.$$ □

Theorem 6.4. *Let A, B be matrices and c a scalar. The following hold (whenever the sizes of A, B allow the expressions to make sense):*

(1) $\|A + B\| \leq \|A\| + \|B\|$,
(2) $\|AB\| \leq \|A\|\|B\|$,
(3) $\|cA\| = |c|\|A\|$.

Proof. (1) For all i, j, we have $|a_{i,j} + b_{i,j}| \leq |a_{i,j}| + |b_{i,j}|$. Add together these inequalities to get $\|A + B\| \leq \|A\| + \|B\|$.

(2) The (i,j) entry of AB is $\sum_k a_{i,k}b_{k,j}$. For each triple i, j, k, the term $|a_{i,k}||b_{k,j}|$ appears in $\|A\|\|B\|$. Furthermore, all of the terms in that product are ≥ 0, so we get

$$\|AB\| = \sum_{i,j} |\sum_k a_{i,k}b_{k,j}| \leq \sum_{i,j} \sum_k |a_{i,k}||b_{k,j}| \leq \|A\|\|B\|.$$

(3) Follows from the fact that $|ca_{i,j}| = |c||a_{i,j}|$. □

Theorem 6.5. *Let* $(C_k)_{k\geq 1}$ *be a sequence of matrices such that* $\sum_{k\geq 1}\|C_k\|$ *converges. Then* $\sum_{k\geq 1} C_k$ *also converges.*

Proof. For each (i,j), we have $\|C_k\| \geq |c_{ij}^{(k)}|$. Recall that a non-negative and bounded sequence converges – the partial sums $\sum_{k=1}^{N}|c_{ij}^{(k)}|$ are non-negative and bounded from above by $\sum_{k\geq 1}\|C_k\|$ and hence converge. This means that $\sum_{k\geq 1} c_{ij}^{(k)}$ is absolutely convergent, and hence convergent. ☐

Corollary 6.6. *Let* A *be an* $n \times n$ *matrix. Then* $\sum_{k\geq 1} A^k/k!$ *converges.*

Proof. We set $C_k = A^k/k!$. Then $\|C_k\| \leq \|A\|^k/k!$. The sum $\sum_{k\geq 1}\|A\|^k/k!$ is convergent (and equal to $e^{\|A\|}-1$), and hence $\sum_{k\geq 1}\|C_k\|$ is also convergent (it is non-negative and bounded from above), so by the previous theorem, we are done. ☐

Now that the sum is well-defined, we can make the definition

$$e^A = I_n + \sum_{k\geq 1}\frac{A^k}{k!}.$$

We will use the convention that $A^0 = I$ for any matrix A, so that the sum can be written more succinctly as

$$e^A = \sum_{k\geq 0}\frac{A^k}{k!}.$$

Theorem 6.7. *If* A *and* B *are* $n\times n$ *matrices which commute, i.e.,* $AB = BA$*, then* $e^{A+B} = e^A e^B$*.*

Proof. e^{A+B} is the limit of the series $\sum_{k\geq 0}\frac{(A+B)^k}{k!}$. Let $\binom{k}{i} = \frac{k!}{i!(k-i)!}$. Since $AB = BA$, we can expand $(A+B)^k$ using the binomial theorem:

$$\frac{(A+B)^k}{k!} = \frac{1}{k!}\sum_{i=0}^{k}\binom{k}{i}A^iB^{k-i} = \sum_{i=0}^{k}\frac{A^i}{i!}\frac{B^{k-i}}{(k-i)!}.$$

If we sum this over $k = 0, 1, \dots,$ we can factor the sum:

$$e^{A+B} = \sum_{k\geq 0}\sum_{i=0}^{k}\frac{A^i}{i!}\frac{B^{k-i}}{(k-i)!} = \left(\sum_{i\geq 0}\frac{A^i}{i!}\right)\left(\sum_{j\geq 0}\frac{B^j}{j!}\right) = e^A e^B.$$ ☐

Corollary 6.8. *For any square matrix* A*, we have* $e^A e^{-A} = I$*. In particular,* e^A *is invertible.*

Proof. A and $-A$ commute with each other, and $e^{A-A} = e^0 = I$. ☐

6.3. **Differential equations satisfied by** e^{tA}**.** Let A be a square matrix. Consider the matrix-valued function

$$E(t) = e^{tA}.$$

Recall that we define the derivative of a matrix-valued function by taking derivatives in each entry separately and this satisfies the product rule.

Theorem 6.9. $E'(t) = AE(t)$.

Proof. Let $c_{ij}^{(k)}$ be the (i,j)-entry of A^k. The (i,j)-entry of $E(t)$ is

$$\sum_{k\ge 0} \frac{c_{ij}^{(k)} t^k}{k!}.$$

Taking the derivative (which can be done term-by-term since the power series converges in an interval, namely everywhere), we get

$$\sum_{k\ge 1} \frac{c_{ij}^{(k)} t^{k-1}}{(k-1)!} = \sum_{k\ge 0} \frac{c_{ij}^{(k+1)} t^k}{k!}.$$

So we get

$$E'(t) = \sum_{k\ge 0} \frac{A^{k+1} t^k}{k!} = AE(t).$$ □

Theorem 6.10. *Let A be an $n \times n$ square matrix and let $B \in \mathbf{R}^n$ be an n-dimensional column vector. Pick $a \in \mathbf{R}$. The differential equation with initial condition*

$$Y'(t) = AY(t), \qquad Y(a) = B$$

has a unique solution on $(-\infty, \infty)$, which is $Y(t) = e^{(t-a)A}B$.

Proof. The derivative of $e^{(t-a)A}B$ is $Ae^{(t-a)A}B$, so it satisfies the differential equation and initial condition, so we just have to show that it is the only solution.

Suppose that $Z(t)$ is another solution, i.e., that $Z'(t) = AZ(t)$ and $Z(a) = B$. Set $G(t) = e^{-(t-a)A}Z(t)$. Then

$$G'(t) = -Ae^{-(t-a)A}Z(t) + e^{-(t-a)A}Z'(t) = -Ae^{-(t-a)A}Z(t) + e^{-(t-a)A}AZ(t) = 0$$

where the last equality follows since $e^{-(t-a)A}A = Ae^{-(t-a)A}$. In particular, $G(t)$ is a constant matrix (independent of t). In particular, $G(t) = G(a) = Z(a) = B$. This implies that $Z(t) = e^{(t-a)A}B = Y(t)$. □

Remark 6.11. Combined with Remark 6.1, this theorem gives a proof of the existence-uniqueness theorem (Theorem 5.1) in the constant-coefficient case. □

We can use this to solve the constant-coefficient non-homogeneous case as well. Consider the equation

$$Y'(t) = AY(t) + Q(t)$$

where A is an $n \times n$ matrix of scalars and $Q(t)$ is a n-dimensional column vector of functions which are continuous on an interval J.

Theorem 6.12. *Notation as above. Pick $a \in J$. The differential equation with initial condition*

$$Y'(t) = AY(t) + Q(t), \qquad Y(a) = B,$$

has a unique solution on J, which is

$$Y(x) = e^{(x-a)A}B + e^{xA}\int_a^x e^{-tA}Q(t)\,\mathrm{d}t.$$

Proof. The derivative of our proposed solution is

$$Ae^{(x-a)A}B + Ae^{xA}\int_a^x e^{-tA}Q(t)\,\mathrm{d}t + e^{xA}e^{-xA}Q(x),$$

so it satisfies the differential equation. Also, setting $x = a$ gives B so it also satisfies the initial condition.

Suppose that $Z(x)$ is another solution to the differential equation. Then $Y(x) - Z(x)$ is a solution to the homogeneous equation (with no $Q(t)$), and $Y(a) - Z(a) = 0$. So by the uniqueness statement in Theorem 6.10, we have $Y(x) - Z(x) = 0$ (as functions), and hence $Y(x) = Z(x)$. $\square$

6.4. **Calculating matrix exponentials for diagonalizable matrices.** If A is diagonal with diagonal entries $\lambda_1, \dots, \lambda_n$, then A^k is also diagonal with entries $\lambda_1^k, \dots, \lambda_n^k$. In that case, the matrix exponential is easy to calculate since we just get a diagonal matrix with entries $e^{\lambda_1}, \dots, e^{\lambda_n}$.

Next, suppose that A is diagonalizable and write it as CDC^{-1} where D is diagonal with entries $\lambda_1, \dots, \lambda_n$. Then $(CDC^{-1})^k = CD^kC^{-1}$, so we get

$$e^A = \sum_{k\ge 0}\frac{A^k}{k!} = \sum_{k\ge 0}\frac{CD^kC^{-1}}{k!} = C\left(\sum_{k\ge 0}\frac{D^k}{k!}\right)C^{-1} = C\mathrm{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})C^{-1}.$$

Example 6.13. Consider the matrix $A = \begin{pmatrix} 6 & -1 \\ 2 & 3 \end{pmatrix}$. We will skip the computation of eigenvalues and eigenvectors; the eigenvalues of A are 5 and 4, and for eigenvectors, we can take $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. So we have

$$A = CDC^{-1}, \qquad D = \begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \qquad C^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

So we get

$$e^A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} e^5 & 0 \\ 0 & e^4 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2e^5 - e^4 & -e^5 + e^4 \\ 2e^5 - 2e^4 & -e^5 + 2e^4 \end{pmatrix}. \qquad \square$$

Example 6.14. The method above will also work if the eigenvalues are complex. Consider $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. The characteristic polynomial is $t^2 + 1$ so its eigenvalues are $i, -i$. For eigenvectors, we can take $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ and $\begin{pmatrix} 1 \\ i \end{pmatrix}$. So we have

$$A = CDC^{-1}, \qquad D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}, \qquad C^{-1} = \frac{1}{2}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}.$$

So we get

$$e^A = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}\begin{pmatrix} e^i & 0 \\ 0 & e^{-i} \end{pmatrix}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} = \begin{pmatrix} \cos(1) & -\sin(1) \\ \sin(1) & \cos(1) \end{pmatrix},$$

where in the last equality, we used the identity $e^{ix} = \cos(x) + i\sin(x)$.

We can also verify this by hand without using complex numbers. The powers of A happen to have nice expressions:

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad A^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad A^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and so, in general, $A^{4+k} = A^k$. We split the sum for e^A into the even and odd powers:

$$e^A = \sum_{k \geq 0} \frac{A^{2k}}{(2k)!} + \sum_{k \geq 0} \frac{A^{2k+1}}{(2k+1)!} = \sum_{k \geq 0} \frac{1}{(2k)!} \begin{pmatrix} (-1)^k & 0 \\ 0 & (-1)^k \end{pmatrix} + \sum_{k \geq 0} \frac{1}{(2k+1)!} \begin{pmatrix} 0 & (-1)^{k+1} \\ (-1)^k & 0 \end{pmatrix}.$$

We can recognize those sums as the Taylor series for $\sin(x)$ and $\cos(x)$ with $x = 1$, so this gives another way to get our answer. □

Example 6.15. Unfortunately, not all matrices are diagonalizable, for example consider $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. Its characteristic polynomial is $(t - \lambda)^2$, but the eigenspace for λ is only 1-dimensional. Fortunately, this one can be done by hand using the law of exponents. Write

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Call these two matrices D and N. Then $DN = ND$, so we get

$$e^A = e^D e^N = \begin{pmatrix} e^\lambda & 0 \\ 0 & e^\lambda \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^\lambda & e^\lambda \\ 0 & e^\lambda \end{pmatrix}$$

where the calculation of e^N follows because $N^2 = 0$. □

Actually, these 3 examples are illustrative of the general situations for 2×2 matrices. Given a 2×2 matrix A, there are 3 possibilities:

(1) A is diagonalizable with real eigenvalues,
(2) A is diagonalizable with complex eigenvalues,
(3) A is not diagonalizable.

The first two cases are handled the same exact way, so we don't need to distinguish them actually. In the third case, the only thing that can go wrong is that the characteristic polynomial of A has a repeated root, so looks like $(t - \lambda)^2$, but the λ-eigenspace only has dimension 1. In general for that case, we can always find an invertible matrix C so that

$$A = C \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} C^{-1}.$$

Remark 6.16. Here's how you find C: first let v be a nonzero eigenvector for A with eigenvalue λ: $Av = \lambda v$. Then find another nonzero vector w satisfying the equation $Aw = \lambda w + v$. Let C be the matrix whose columns are v and w (in that order). We won't go into the details of why this is possible or do these kinds of computations. □

For general $n \times n$ matrices A, one can appeal to the existence of **Jordan canonical form**. This says that we can find an invertible matrix C (we have to allow the possibility of complex entries) so that

$$CAC^{-1} = D + N$$

where D is diagonal (possibly complex) and N is nilpotent (i.e., $N^n = 0$) and $DN = ND$. In fact, we can also find these matrices so that N is 0 everywhere except on the superdiagonal (the entries in row i and column $i+1$ for some i) where the entries are either 0 or 1. We won't discuss Jordan canonical form here, you'll see it in Math 542. But once you have it, you can compute e^A using the ideas from the examples above.

There are some other methods for calculating e^A in §7.13 and §7.14 in Apostol, but we won't discuss these.

6.5. **Proof of uniqueness-existence for linear systems of differential equations.** We now return to our general system of equations

$$Y' = P(t)Y + Q(t), \qquad Y(a) = B$$

where P, Q are continuous on some interval J and $a \in J$. We have already discussed how to solve this when $P(t)$ is constant and we have a uniqueness-existence theorem for the solution. Now we discuss this more general situation. A method for obtaining a solution is discussed in §7.18 of Apostol. We will not go into the details; the formula is rather complicated and I think it's a better use of time to discuss the proof of uniqueness-existence since it will involve some mathematical ideas you likely have not come across yet.

Here is the formal statement:

Theorem 6.17. *Let J be an (open) interval. Let $A(t)$ be an $n \times n$ matrix of functions which are continuous on J. Pick $a \in J$ and let $B \in \mathbf{R}^n$ be a column vector. Then the system of differential equations with initial condition*

$$Y'(t) = A(t)Y(t), \qquad Y(a) = B$$

has exactly one solution on J.

First, we have to show that some solution exists, and we will produce one using Picard's method of successive approximations. The idea is to start with some proposed solution, and try to successively improve it. For example, first start with the constant $Y_0(t) = B$. This generally won't be a solution unless $A(t)$ is identically 0. In any case, we substitute Y_0 into Y in the right hand side of the equation to get

$$Y'(t) = A(t)B, \qquad Y(0) = B$$

However the left hand side does not involve Y anymore, so this is something easy to solve using first-year calculus, the answer is

$$Y_1(x) = B + \int_a^x A(t)B\,\mathrm{d}t.$$

Now we substitute Y_1 into the right hand side to get

$$Y'(t) = A(t)Y_1(t), \qquad Y(0) = B,$$

which we can again solve:

$$Y_2(x) = B + \int_a^x A(t)Y_1(t)\,\mathrm{d}t,$$

and so on... so we define $Y_{i+1}(x) = B + \int_a^x A(t)Y_i(t)\,\mathrm{d}t$ in general. Here we are thinking of the sequence $Y_0, Y_1, Y_2, \ldots$ as approximations of the actual solution which get "closer" as we do more steps.

Example 6.18. Consider the constant-coefficient case $Y'(t) = AY(t)$ with $Y(0) = B$. Take $J = (-\infty, \infty)$ and $a = 0$. If we use the idea above, our functions are

$$\begin{aligned} Y_0(t) &= B, \\ Y_1(t) &= B + \int_0^x AB\,\mathrm{d}t = B + xAB, \\ Y_2(t) &= B + \int_0^x A(B + tAB)\,\mathrm{d}t = B + xAB + \frac{x^2}{2}A^2B, \\ Y_3(t) &= B + \int_0^x A(B + xAB + \frac{x^2}{2}A^2B)\,\mathrm{d}t = B + xAB + \frac{x^2}{2}A^2B + \frac{x^3}{6}A^3B, \end{aligned}$$

and in general we would get $Y_m(t) = \sum_{i=0}^m \frac{x^i}{i!}A^iB$. Taking the limit $m \to \infty$, we get $Y(t) = e^A B$, which we've already seen is the solution. □

Fortunately, this works in general! Here's the statement:

Theorem 6.19. *Notation as above. Define a function $Y \colon J \to \mathbf{R}$ by*

$$Y(x) = \lim_{k\to\infty} Y_k(x)$$

for $x \in J$. Then

1. *For all $x \in J$, the limit above exists.*
2. *Y is a continuous function.*
3. *$Y(x) = B + \int_a^x A(t)Y(t)\,\mathrm{d}t$ for all $x \in J$. In particular, $Y(a) = B$.*
4. *Y is differentiable, and $Y'(x) = A(x)Y(x)$ for all $x \in J$.*

Proof. (1) First rewrite Y_k as a telescoping sum:

$$Y_k(x) = Y_0(x) + \sum_{m=0}^{k-1}(Y_{m+1}(x) - Y_m(x)).$$

It is enough to show that the telescoping sum converges for all x when $k \to \infty$. To do that, we will show that the infinite sum of norms

$$\sum_{m\ge 0} \|Y_{m+1}(x) - Y_m(x)\|$$

converges for all x. So pick a specific value of x. The functions in the entries of $A(t)$ are continuous on the closed interval whose endpoints are a to x, so they are bounded. Let M be a value which bounds the sum of all of these functions on this closed interval. Then we claim that

$$\|Y_{m+1}(x) - Y_m(x)\| \le \|B\|\frac{M^{m+1}|x-a|^{m+1}}{(m+1)!}.$$

We will prove this by induction on m. When $m = 0$, we have

$$\|Y_1(x) - Y_0(x)\| = \left\|\int_a^x A(t)B\,\mathrm{d}t\right\| \le \pm\int_a^x \|A(t)\|\|B\|\,\mathrm{d}t \le \pm\int_a^x M\|B\|\,\mathrm{d}t \le \|B\|M|x-a|,$$

where we use that the absolute value of an integral is bounded from above by the integral of the absolute value and the sign is $+$ if $a < x$ and $-$ if $a > x$.

For the induction step with $m > 0$, we have

$$\begin{aligned}\|Y_{m+1}(x) - Y_m(x)\| &= \left\| \int_a^x A(t)(Y_m(t) - Y_{m-1}(t))\,\mathrm{d}t \right\| \\ &\leq \pm \int_a^x \|A(t)\|\|Y_m(t) - Y_{m-1}(t)\|\,\mathrm{d}t \\ &\leq \pm \int_a^x M\|B\| \frac{M^m|t-a|^m}{m!}\,\mathrm{d}t \\ &\leq \|B\| \frac{M^{m+1}|x-a|^{m+1}}{(m+1)!},\end{aligned}$$

where in the third line we used the induction hypothesis and again the sign is $+$ if $a < x$ and $-$ if $a > x$. This proves the induction step, so the claim is proven.

Next, we have

$$\sum_{m \geq 0} \|Y_{m+1}(x) - Y_m(x)\| \leq \|B\| \sum_{m \geq 0} \frac{(M|x-a|)^{m+1}}{(m+1)!} = \|B\|(e^{M|x-a|} - 1),$$

so the sum on the left converges since it has all non-negative terms and is bounded from above.

(2) We will skip the proof of (2). This is discussed in §7.21 in Apostol, but relies on the notion of *uniform convergence* which we have not discussed, and is covered in a later course (Math 521).

(3) Consider the equations

$$Y_{k+1}(x) = B + \int_a^x A(t)Y_k(t)\,\mathrm{d}t.$$

Take the limit of both sides and exchange the integral with the limit (this requires an appeal to uniform convergence to justify, so we omit it) to get

$$Y(x) = B + \int_a^x A(t)Y(t)\,\mathrm{d}t.$$

If we take $x = a$, the right side just becomes B, so $Y(a) = B$.

(4) Since Y is continuous on J, the integrand $A(t)Y(t)$ is also continuous on J, so the function $\int_a^x A(t)Y(t)\,\mathrm{d}t$ is differentiable by the first fundamental theorem of calculus (applied to each term of the matrix). Adding constants from B doesn't affect that. Finally, taking derivatives then yields $Y'(x) = A(x)Y(x)$. □

Proof of Theorem 6.17. We have just constructed a solution, so it remains to show that there can only be one solution. Let $Y(x)$ and $Z(x)$ both be solutions to the differential equation with initial condition.

Using the differential equation, we have

$$Z'(t) - Y'(t) = A(t)(Z(t) - Y(t)).$$

Integrate this from a to x to get

$$Z(x) - Z(a) - Y(x) + Y(a) = \int_a^x A(t)(Z(t) - Y(t))\,\mathrm{d}t.$$

Since $Z(a) = B = Y(a)$, those terms cancel on the left side. As before, continuous functions on the closed interval with endpoints x and a are bounded. Let M be a bound for $\|A(t)\|$ and let M_1 be a bound for $\|Z(t) - Y(t)\|$ on this interval. We claim that

$$\|Z(x) - Y(x)\| \le M^m M_1 \frac{|x-a|^m}{m!}$$

for all integers $m \ge 0$. We prove this by induction. The base case $m = 0$ follows from the definition of M_1. Now we do the induction step and assume it holds for some $m > 0$. Then we get

$$\begin{aligned} \|Z(x) - Y(x)\| &\le \pm \int_a^x \|A(t)\| \|Z(t) - Y(t)\| \, \mathrm{d}t \\ &\le \pm \int_a^x M(M^m M_1 \frac{|t-a|^m}{m!}) \, \mathrm{d}t \\ &\le M^{m+1} M_1 \frac{|x-a|^{m+1}}{(m+1)!}. \end{aligned}$$

where the sign is $+$ if $a < x$ and $-$ if $a > x$. This proves the induction step, and hence the general claim.

Now take the limit as $m \to \infty$ of both sides of the inequality:

$$\|Z(x) - Y(x)\| \le \lim_{m\to\infty} M_1 \frac{(M|x-a|)^m}{m!}.$$

The limit on the right is 0 (this is a general fact from first-year calculus: $\lim_{m\to\infty} |x|^m/m! = 0$ for any x), so $\|Z(x) - Y(x)\| = 0$, which is only possible if $Z(x) = Y(x)$. □

Finally, we can consider the case when $Q(t)$ is also present. The existence of a solution is given in §7.18 of Apostol, and uniqueness follows from uniqueness in the homogeneous case since the difference of any two solutions of the non-homogeneous equation is a solution of the homogeneous one with the same initial condition as the 0 solution.

Even if this method is impractical to carry out by hand, we see that it has an important theoretical consequence. However, it can also be used for actual approxiations. Say we only want to approximate the values of $Y(x)$ at some number x. Then the steps above could be performed some number of times (the integrals being approximated as well) and we'd get an approximate solution. To understand how good this approximation actually is, we would need to study the error terms, which is beyond the scope of what we will discuss.

6.6. **Non-linear first-order systems.** Thus far, we've only talked about linear differential equations, i.e., they only involve derivatives of our functions, the original functions, and linear combinations (possibly with functions as our coefficients), but we never multiply them together. An example of a non-linear differential equation is

$$y'(t) = t^2 + y(t)^2.$$

We won't really expect to be able to find explicit solutions in general, but it would at the very least be useful to have uniqueness-existence theorems for these kinds of equations. We restrict to first-order, meaning we don't take the derivative more than once. However, we'll continue to allow matrix-valued functions. So our general setup is

$$Y'(t) = F(t, Y(t)), \qquad Y(a) = B$$

where Y is an n-dimensional column vector of functions, and F is also an n-dimensional column vector of functions. This generalizes our previous situation if we take $F(t, Y(t)) = A(t)Y(t)$. We can try to use Picard's method of successive approximations in this situation as well. So we'd define a sequence of column vectors by

$$Y_0(x) = B, \qquad Y_{k+1}(x) = B + \int_a^x F(t, Y_k(t))\,\mathrm{d}t \quad (k \geq 0).$$

This won't generally work (the proof before required that we could bound certain terms), but it will work if we impose some extra conditions.

Again we have a uniqueness-existence theorem (known as the Picard–Lindelöf theorem) in the presence of some hypotheses. They are a bit technical, so I will point you to §7.23 of Apostol for the statement. The precise details aren't going to be super important for us in this class. What I want to discuss instead is a more general context that this fits into.

6.7. **Contractions and Banach fixed-point theorem.** Go back to the differential equation

$$Y'(t) = A(t)Y(t), \qquad Y(a) = B.$$

Given a continuous function $f\colon \mathbf{R} \to \mathbf{R}^n$, we defined a new function $T(f)\colon \mathbf{R} \to \mathbf{R}^n$ by $T(f)(x) = B + \int_a^x A(t)Y(t)\,\mathrm{d}t$ and repeatedly applied the transformation T to an initial guess $Y_0(x) = B$. So we obtained a sequence $Y_0, T(Y_0), T^2(Y_0), T^3(Y_0), \dots$ and showed that the limit $\lim_{k\to\infty} T^k(Y_0)$ exists, and is a solution to the original equation. To put this into a more general framework, we make some definitions.

Definition 6.20. Let V be a real vector space. A function $N\colon V \to \mathbf{R}$ is a **norm** if it satisfies the following properties:

(1) $N(x) \geq 0$ for all $x \in V$,
(2) $N(cx) = |c|N(x)$ for all $c \in \mathbf{R}$ and $x \in V$,
(3) (Triangle inequality) $N(x + y) \leq N(x) + N(y)$ for all $x, y \in V$,
(4) $N(x) = 0$ implies $x = 0$.

A **normed vector space** is a pair (V, N) where V is a vector space and N is a norm on it. Usually we just write V instead of (V, N). □

Instead of $N(x)$, we will usually just write $\|x\|$. We usually think of $\|x\|$ as the "size" of x, and so $\|x - y\|$ can be thought of as the "distance" between x and y.

Example 6.21. (1) $V = \mathbf{R}^n$ and $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$. This is the Euclidean norm.
(2) $V = \mathbf{R}^n$ and $\|x\|_1 = |x_1| + \cdots + |x_n|$. When V is the space of matrices of some size, this is the matrix norm we discussed earlier.
(3) Choose real numbers $a < b$ and let $\mathscr{C}([a, b])$ be the vector space of continuous functions $\varphi\colon [a, b] \to \mathbf{R}$. The max norm is $\|\varphi\| = \max_{a \leq x \leq b} |\varphi(x)|$. More generally, we can consider the space of functions from $[a, b]$ to $\mathbf{R}^n$ and define the max norm by $\|\varphi\| = \max_{a \leq x \leq b} \|\varphi(x)\|_1$. □

Definition 6.22. Let V be a normed vector space.

(1) A sequence $x_1, x_2, \dots \in V$ is a **Cauchy sequence** if for all $\varepsilon > 0$, there exists N so that $n, m \geq N$ implies that $\|x_n - x_m\| < \varepsilon$.
(2) A sequence $x_1, x_2, \dots \in V$ **converges** to $x \in V$ if $\lim_{n\to\infty} \|x - x_n\| = 0$. □

Intuitively, a sequence is Cauchy if the points are getting very close to each other as we go farther out into the sequence. Using the triangle inequality, we can show that any sequence that converges to some value must be a Cauchy sequence. Furthermore, the triangle inequality can be used that a sequence has a unique limit, meaning there it cannot converge to two different points. There are examples of Cauchy sequences that do not converge, but they will in examples of interest.

Definition 6.23. A normed vector space is **complete** if every Cauchy sequence converges to some value. A **Banach space** is a complete normed vector space. $\square$

The examples of norms mentioned above are all complete, so they all define Banach spaces. The fact that the max norm is complete is a non-trivial statement which is something we won't discuss. While it is not bad to show that a Cauchy sequence under max norm has a limit to some function, the hard part is to prove that this resulting function is still continuous.

Definition 6.24. Let V, W be Banach spaces. A function $f: V \to W$ is a **contraction** if there exists a constant c with $0 \le c < 1$ such that

$$c\|x - y\|_V \ge \|f(x) - f(y)\|_W$$

for all $x, y \in V$. (Here the subscript denotes where the norms are computed.) $\square$

In other words, a contraction is a function which shrinks distances between points (but it is important that it shrinks by a guarenteed ratio, which is the constant c).

Theorem 6.25 (Banach fixed point theorem)**.** *Let V be a Banach space. Let $f: V \to V$ be a contraction. Then there exists a unique $x \in V$ such that $f(x) = x$.*

This can be proven in a more general context where V only has the notion of distance but is not necessarily a vector space ("metric spaces"). We will see how this applies to existence-uniqueness theorems by considering the case when V is the space of continuous matrix-valued functions on $[a, b]$.

Proof. Let x_0 be any point in V, and in general, define $x_n = f(x_{n-1}) = f^n(x_0)$. We claim that $x_0, x_1, x_2, \dots$ is a Cauchy sequence. Let c be the contraction constant for f. First, we have

$$\|x_n - x_{n-1}\| \le c\|x_{n-1} - x_{n-2}\| \le c^2\|x_{n-2} - x_{n-3}\| \le \cdots \le c^{n-1}\|x_1 - x_0\|.$$

By repeatedly using the triangle inequality, for $n \ge m$, we have

$$\|x_n - x_m\| \le \|x_n - x_{n-1}\| + \|x_{n-1} - x_{n-2}\| + \cdots + \|x_{m+1} - x_m\|.$$

Combining these two inequalities, we get

$$\|x_n - x_m\| \le (c^{n-1} + c^{n-2} + \cdots + c^m)\|x_1 - x_0\| = c^m \frac{1 - c^{n-m}}{1 - c}\|x_1 - x_0\| \le c^m \frac{\|x_1 - x_0\|}{1 - c}.$$

The point is that the last expression can be made arbitrarily small by taking m large enough since $c < 1$. Hence this is a Cauchy sequences, so $\lim_{n\to\infty} x_n$ exists by completeness, call the limit x^*. This means that $\lim_{n\to\infty} \|x^* - x_n\| = 0$. Since f is a contraction, we have $\lim_{n\to\infty} \|f(x^*) - f(x_n)\| \le c \lim_{n\to\infty} \|x^* - x_n\| = 0$, so $f(x^*) = \lim_{n\to\infty} f(x_n)$. But this is the same sequence as before (with x_0 missing), so $f(x^*) = x^*$.

Finally, suppose there is another point z such that $f(z) = z$. Then $\|z - x^*\| \le c\|z - x^*\|$ which implies that $\|z - x^*\| = 0$ (otherwise we could divide and get $1 \le c$). But this means $z = x^*$. $\square$

6.8. **Applications to differential equations.** Let V be the space of continuous functions $\varphi\colon [a,b] \to \mathbf{R}^n$ with the max norm

$$\|\varphi\| = \max_{a \le x \le b}(|\varphi_1(x)| + \cdots + |\varphi_n(x)|).$$

We have the differential equation

$$Y'(t) = A(t)Y(t), \qquad Y(\alpha) = B$$

where $a < \alpha < b$. Define a function

$$T\colon V \to V$$
$$T(\varphi)(x) = B + \int_\alpha^x A(t)\varphi(t)\,\mathrm{d}t.$$

As in the proof of Theorem 6.19, the entries of the function $A(t)$ are bounded on $[a,b]$, so set $M = \max_{a \le t \le b} \|A(t)\|$ (here we take max over matrix norm). Then for two functions φ, ψ, we have

$$(T(\varphi) - T(\psi))(x) = \int_\alpha^x A(t)(\varphi(t) - \psi(t))\,\mathrm{d}t,$$

and so the maximum over all x is then

$$\|T(\varphi) - T(\psi)\| \le |b-a|M\|\varphi - \psi\|.$$

If $|b-a|M < 1$, then T is a contraction. If we make the interval smaller, M will still be an upper bound for $\max_t \|A(t)\|$. So to force T to be a contraction, we can pick a closed interval around α which has length $< 1/M$.

In that case, the fixed point theorem tells us that there is a unique continuous function Y (on this smaller interval) such that $T(Y) = Y$. This means that

$$Y(x) = B + \int_\alpha^x A(t)Y(t)\,\mathrm{d}t.$$

In particular, $Y(\alpha) = B$, and taking derivatives shows that $Y'(x) = A(x)Y(x)$, which is what we wanted. Of course, we want a solution not just on the small interval. To get around this, we can pick two different intervals whose lengths are $< 1/M$ which contain α. Suppose they both contain an open interval around α. Then the functions produced by the fixed-point theorem must agree on this open interval by uniqueness. Hence, we end up with a function defined on the union of these two intervals. We can keep extending our solution by changing α to a different value in the union of these two intervals. In this way, we can guarantee a solution on the whole interval $[a,b]$.

Looking at the more general situation

$$Y'(t) = F(t, Y(t)), \qquad Y(\alpha) = B$$

on some interval $[a,b]$, we can mimick the situation and define $T\colon V \to V$ by

$$T(\varphi)(x) = B + \int_\alpha^x F(t, \varphi(t)\,\mathrm{d}t.$$

If we want this to be a contraction, we can restrict our attention to functions F that satisfy the additional condition

$$\|F(t,\varphi) - F(t,\psi)\| \le c\|\varphi - \psi\|$$

for some constant c. Then, as above, we can show that

$$\|T(\varphi) - T(\psi)\| \le |b-a|c\|\varphi - \psi\|$$

and shrink our interval so that its length is $< 1/c$. The fixed point is then the unique solution on this small interval, and we can extend it in the same way as discussed above.