SPP Analysis Supplementary Notes

Zheng-Chao Han
zchan@math.rutgers.edu

January 14, 2022

1 Introductory Remarks

The goals of the pre-enrollment "Boot Camp" were to provide an in-depth review and to fill in gaps in some background material expected in standard first year graduate courses (mostly on topics in abstract linear algebra and advanced calculus which have often not be adequately treated in undergraduate courses). Problem sessions are integral parts of this program. Problems to be discussed will emphasize interconnectedness between different areas of mathematics, and will include some problems from the written qualifying exams. Attention will be paid to proper writing of mathematical proofs; critiques of student solutions will be provided.

Topics related to advanced calculus include:

- Basic properties of the reals: Limits (including upper and lower limits), Cauchy sequences, completeness, sequential compactness (Bolzano–Weierstrass theorem) and compactness (Heine-Borel Theorem).
- Basic tools: Cauchy-Schwarz inequality. Summation (integration) by parts.
- Sequences and series of numbers and functions, including absolute and uniform convergence, and equicontinuity. Applications involving power series, integration and differentiation.
- Basic topological notions such as connectivity, Hausdorff spaces, compactness, product spaces and quotient spaces. Emphasis on examples in Euclidean and metric spaces.
- Compactness criteria in metric spaces. Arzelà–Ascoli Theorem and applications.
- Review of multiple, line and surface integrals, theorems of Green and Stokes and the divergence theorem.
- Jacobians, implicit and inverse function theorems, and applications. Change of variables formula. Role of exterior calculus.

Almost every incoming student probably has seen the majority of the topics above. What prepares a student do well in the first year graduate courses is not just exposure to these topics, but depth of understanding and fluency of applying them.

We will spend minimum amount of time reviewing the relevant topics formally; we will spend more time on sets of problems to explain the intuition behind the concepts and methods, how they are applied, and help students to gain a better understanding of these concepts and methods.

Professors Ocone and Mirek have done these analysis sessions in the last couple years, and each has prepared some very nice written notes. I will share those notes and rely on them. Their notes have somewhat different emphasis and style, with Professor Ocone's notes being more basic and aligned with the topics listed above, while Professor Mirek's notes going more into some generalities and depth, with some topics typically on the graduate analysis curriculum and with more examples rooted in number theory.

There is not enough allotted time to go over the material in either notes. Since the incoming students have varying backgrounds and interests, these notes give students enough flexibility to find material suitable to their background and interest and work on those parts in a more focused way.

I am also preparing some supplementary notes and comments. The notes by Professors Ocone and Mirek contain some summary reviews of the topics discussed, and some sketch of proofs. My focus in these notes will not be on a review of topics, but on strategy of problem solving. In order to make our discussions more productive, I ask that students do some initial study prior to the start of our boot camp, write up solutions for as many problems in the posted diagnostic quiz as you have time for before the boot camp starts, and submit them and any additional topics or problems that you would like us to discuss in our sessions.

1.1 Some General Guidance

Here are a few points to keep in mind when working on a problem.

- Learn to identify some key steps which may lead to a solution of the problem and how the given information may help to carry out these steps.
- Learn to keep focus on the key steps and set aside some non-essential technicalities at the beginning.
- Learn to carry out some reductions to reduce the problem to one or more simpler sub-problems, and start with concrete/simpler cases to test out how your ideas may or may not work out.

One can see that these suggestions are often different from one's typical learning experience in a math course, which is often going from the abstract to the concrete, from the general to the specific, and, to use a variation of an often quoted saying, "carrying a hammer to look for a nail to pound". A good strategy to solve a problem is to start with the problem and *identify its key features and find or create the right tools which may bring about a solution.*

The suggestions above apply for mathematical problem solving in general. In analysis, one needs to learn to "size up" different quantities and identify the ones which play a leading order role in solving our problem. We begin with a few examples to illustrate the suggested strategies discussed above.

1.2 Some Initial Examples

Example 1.1 Assume that $x_n \to A$ as $n \to \infty$. Prove that $\frac{x_1+\cdots+x_n}{n} \to A$ as $n \to \infty$.

We first do an analysis of the problem. From the given assumption, for all sufficiently large n, say, for $n > N$, x_n is within $\epsilon > 0$ of A. How would this indicate that $\frac{x_1+\cdots+x_n}{n}$ is close to A for large n?

We examine

$$\frac{x_1 + \cdots + x_n}{n} = \frac{x_1 + \cdots + x_N + x_{N+1} + \cdots + x_n}{n},$$

in which all terms $x_{N+1}, \cdots, x_n$ are within $\epsilon > 0$ of A. How do we handle the first N terms? Well, when n is sufficiently large, the effect of $\frac{x_1+\cdots+x_N}{n}$ is negligible.

So we can see a solution strategy now. But we can do one small reduction: set $y_n = x_n - A$. Then $y_n \to 0$ and $\frac{x_1+\cdots+x_n}{n} = A + \frac{y_1+\cdots+y_n}{n}$, so it suffices to prove $\frac{y_1+\cdots+y_n}{n} \to 0$. We will leave it to the student to write up a proof for this last statement using the strategy discussed above.

Checkpoint 1.2 Find the limit $\lim_{n\to\infty}\sqrt[n]{x_1\cdots x_n}$ if $x_n > 0$ and $x_n \to A > 0$.

Hint. Use $x = e^{\ln x}$ for $x > 0$.

Example 1.3 Find the limit $\lim_{n\to\infty}\sqrt[n]{\int_a^b |f(x)|^n\,dx}$ for $f \in C[a,b]$.

Here we are dealing with a generic $f \in C[a,b]$. Obviously we can't hope to evaluate the integral $\int_a^b |f(x)|^n\,dx$ except when $|f(x)|$ is a constant c over $[a,b]$, in which case the root of the integral becomes $\sqrt[n]{(b-a)c^n} \to c$. At least in this special case we know the answer.

For the general case, when $|f(x)|$ is not a constant, our focus should be on what affects this integral when $n \to \infty$. When $|f(x_1)| < |f(x_2)|$, $|f(x_1)|^n$ would be significantly smaller than $|f(x_2)|^n$ as $n \to \infty$, so this suggests that the largest contribution would come from those x such that $|f(x)| = M := \max_{[a,b]} |f(y)|$.

First, since $|f(x)| \le M$ for all $x \in [a,b]$, it follows that

$$\sqrt[n]{\int_a^b |f(x)|^n\,dx} \le \sqrt[n]{(b-a)M^n}.$$

Since $\sqrt[n]{(b-a)M^n} \to M$, we can conclude that $\limsup_{n\to\infty}\sqrt[n]{\int_a^b |f(x)|^n\,dx} \le M$.

How does the set of points x such that $|f(x)| = M$ contribute to the integral?

Let x^* be a point such that $|f(x^*)| = M$. Then near x^*, $|f(x)|$ will be close to M. More precisely, for any given $\epsilon > 0$, there must exist some interval I containing x^* with length $\delta > 0$, in which $|f(x)| \ge M - \epsilon$. Then

$$\sqrt[n]{\int_a^b |f(x)|^n\,dx} \ge \sqrt[n]{\int_I (M-\epsilon)^n\,dx} = (M-\epsilon)\sqrt[n]{\delta}.$$

Since $\sqrt[n]{\delta} \to 1$, we conclude that $\liminf_{n\to\infty}\sqrt[n]{\int_a^b |f(x)|^n\,dx} \ge (M-\epsilon)$. Since this inequality holds for any $\epsilon > 0$, we conclude that $\liminf_{n\to\infty}\sqrt[n]{\int_a^b |f(x)|^n\,dx} \ge M$. Combining the two inequalities, we conclude that $\lim_{n\to\infty}\sqrt[n]{\int_a^b |f(x)|^n\,dx}$ exists and equals M.

Checkpoint 1.4 Drills on $\liminf_{n\to\infty}$ and $\limsup_{n\to\infty}$. Here are some often used properties on $\liminf_{n\to\infty}$ and $\limsup_{n\to\infty}$.

1. Properties involving sum of sequences.

$$\liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n$$

$$
\begin{aligned}
&\leq \liminf_{n\to\infty}(a_n+b_n)\\
&\leq \limsup_{n\to\infty} a_n + \liminf_{n\to\infty} b_n\\
&\leq \limsup_{n\to\infty}(a_n+b_n)\\
&\leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n
\end{aligned}
$$

2. $\limsup_{n\to\infty}\sqrt[n]{|a_n|} \leq \limsup_{n\to\infty}|\frac{a_{n+1}}{a_n}|$.

3. $\liminf_{n\to\infty}|\frac{a_{n+1}}{a_n}| \leq \liminf_{n\to\infty}\sqrt[n]{|a_n|}$.

4. If $\lim_{n\to\infty}|\frac{a_{n+1}}{a_n}|$ exists, then $\lim_{n\to\infty}\sqrt[n]{|a_n|}$ exists, and equals $\lim_{n\to\infty}|\frac{a_{n+1}}{a_n}|$. (The converse is not true. Give an example.)

Use the definitions of $\liminf_{n\to\infty}$ and $\limsup_{n\to\infty}$ closely. For example, to prove $\liminf_{n\to\infty} c_n \leq B$ it suffices to prove that for any $\epsilon > 0$, there exists a *subsequence*c_{n_k} such that $c_{n_k} \leq B+\epsilon$ for all sufficiently large k; to prove that $\liminf_{n\to\infty} c_n \geq A$ it suffices to prove that for any $\epsilon > 0$, $c_n \geq A-\epsilon$ for all sufficiently large n (of the *full sequence*).

In analysis we often need to assess the effect of algebraic manipulations of several quantities which are tending to infinity or zero. In carrying out such an analysis it is important to identify the leading order ones and make use of them effectively. The following examples and exercises illustrate how to carry out such an analysis in analyzing algebraic expressions of the form $\infty-\infty, \infty\cdot 0, 0/0$.

Example 1.5 Analyze the limit of $\sqrt{x^2+x+1} - \sqrt[3]{x^3+x+1}$ as $x\to\infty$.

First, we do a rough analysis of the orders of magnitude of the two expressions. When $x\to\infty$, the leading order of $\sqrt{x^2+x+1}$ is $\sqrt{x^2}=|x|$, and the the leading order of $\sqrt[3]{x^3+x+1}$ is $\sqrt[3]{x^3}=|x|$. Thus we need to do a more refined analysis.

One effective way to carry this out is to factor out the leading order terms, and do a Taylor expansion of the factored-out terms, which are now evaluated near a finite value.

$$\sqrt{x^2+x+1} = x\sqrt{1+x^{-1}+x^{-2}} = x\left\{1+\frac{1}{2}\left(x^{-1}+x^{-2}\right)+O(x^{-2})\right\},$$

$$\sqrt[3]{x^3+x+1} = x\sqrt[3]{1+x^{-2}+x^{-3}} = x\ \left\{1+\frac{1}{3}\left(x^{-2}+x^{-3}\right)+O\left(x^{-4}\right)\right\}.$$

In the above, we did a Taylor expansion of $\sqrt{1+u}$ and respectively $\sqrt[3]{1+u}$ at $u=0$, and treated $u=x^{-1}+x^{-2}$ and respectively $u=x^{-2}+x^{-3}$. Thus we have

$$
\begin{aligned}
&\sqrt{x^2+x+1} - \sqrt[3]{x^3+x+1}\\
=&\frac{1}{2}+\frac{1}{2x}+O(x^{-1})-\frac{1}{3x}-\frac{1}{3x^2}-O(x^{-3})\\
\to&\frac{1}{2} \quad \text{as } x\to\infty.
\end{aligned}
$$

Checkpoint 1.6 Find $\lim_{x\to 0} \frac{\sin^{100} x - x^{100}}{x^a}$, where $a > 0$ is some given parameter. This and the following problem were originally used by Professor Kasper Larsen.

l'Hospital's rule is often a student's first choice for dealing with limits of the form 0/0. But that is only a mechanical way of dealing with such a limit. The key is to compare the order of vanishing of the numerator and the denominator.

Hint. Use the Taylor expansion for $\sin x$ in the form of $\sin x = x + R(x)x^3$ where $R(x) \to -\frac{1}{6}$ as $x \to 0$, then use the binormal expansion on $x + R(x)x^3$.

Checkpoint 1.7 Analyze the dependence of $a(t) = \sum_{n=1}^{\infty} e^{-n^p t}$ on t. Define $a(t) = \sum_{n=1}^{\infty} e^{-n^p t}$, where $p > 0$ is some given parameter. Show that there exists some $\alpha > 0$ such that

$$\lim_{t\to 0} \frac{a(t)}{t^\alpha} \text{ exists.}$$

Also identify this limit, and identify conditions on p such that $\int_0^1 a(t)\,dt$ converges. Under what conditions is $\int_1^\infty a(t)\,dt$ convergent?

Hint. For the beginning part, use the integral estimate

$$\int_1^\infty e^{-x^p t}\,dx \le \sum_{n=1}^{\infty} e^{-n^p t} \le \int_0^\infty e^{-x^p t}\,dx$$

and make a change of variables in the integrals to delineate the role of t.

Example 1.8 (January 2017 WQ).

Given any polynomials $z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0$ and $a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$. Evaluate

$$\int_{|z|=r} \frac{a_{n-1}z^{n-1} + \cdots + a_1 z + a_0}{z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0} dz$$

for all sufficiently large $r > 0$.

Even though this problem appeared as a problem in complex analysis, the main insight is still to quantify the size of the quotient $\frac{a_{n-1}z^{n-1}+\cdots+a_1z+a_0}{z^n+b_{n-1}z^{n-1}+\cdots+b_1z+b_0}$ when $|z| = r$ is large. Based on our discussion earlier, we expect it to be approximated by $\frac{a_{n-1}z^{n-1}}{z^n} = \frac{a_{n-1}}{z}$, and the integral of the latter is easy to evaluate.

To carry out this approach rigorously, we need to estimate how close the approximation is as $|z| \to \infty$. One can see that

$$\begin{aligned}
&\left|\frac{a_{n-1}z^{n-1} + \cdots + a_1 z + a_0}{z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0} - \frac{a_{n-1}}{z}\right| \\
=&\left|\frac{a_{n-1}z^n + \cdots + a_1 z^2 + a_0 z - a_{n-1}\left(z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0\right)}{z^{n+1} + b_{n-1}z^n + \cdots + b_1 z^2 + b_0 z}\right| \\
\le&\frac{C}{|z|^2} \quad \text{for all sufficiently large} |z|.
\end{aligned}$$

Since the circle $\{z : |z| = r\}$ has length $2\pi r$, it is now clear that the approximate integral differs from the given integral by an error which tends to 0 as $|z| \to \infty$.

Example 1.9 The first encounter with the Cauchy-Schwarz inequality.

The Cauchy-Schwarz inequality states that, for any $\mathbf{a} = (a_1, \cdots, a_n), \mathbf{b} = (b_1, \ldots, b_n)$,

$$\Big|\sum_{i=1}^{n} a_i b_i\Big| \le \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \left(\sum_{i=1}^{n} b_i^2\right)^{1/2}. \tag{1.1}$$

There are many proofs of this inequality, including very conceptual and short ones. But we will use this occasion to illustrate how one would go about investigating such a problem in the absence of knowing the inner product concept which underlies this inequality.

There are $2n$ quantities in this inequality. One interpretation of this inequality is that, among those values of these $2n$ quantities which make the right hand side a constant, the left hand side can't be made large than the right hand side. But how can we make the left hand side as large as possible? Can we examine the relations between the two sides by varying only one or two variables at a time? Why don't we study the small n cases first, and see whether we can gain some insight there?

Small n cases. The two sides are equal when $n = 1$. The $n = 2$ case takes the form of

$$|a_1 b_1 + a_2 b_2| \le \sqrt{a_1^2 + a_2^2}\sqrt{b_1^2 + b_2^2}. \tag{1.2}$$

There are several possible approaches at this point.

(i). Treat only one quantity, say, a_1, as a variable, and the rest as parameters, and prove that the left hand side is no more than the right hand side as a function of this variable. Algebraically it is easier to try to prove

$$f(a_1) := |a_1 b_1 + a_2 b_2|^2 - (a_1^2 + a_2^2)(b_1^2 + b_2^2) \le 0.$$

One can use one-variable calculus to tackle this. You should carry out the calculus to find out the roles played by the remaining variables. Can you handle the general n case this way too?

If one works directly with the left hand side as set up in (1.2), or try to prove $a_1 b_1 + a_2 b_2 - \sqrt{a_1^2 + a_2^2}\sqrt{b_1^2 + b_2^2} \le 0$, then the calculus is a bit more tedious to handle.

(ii). Formulate the problem as a constrained maximization problem and tackle it using the method of **Lagrange multipliers**. Note that if we change (a_1, a_2) to $t(a_1, a_2)$ for some $t > 0$, then both sides get multiplied by t. We say that they have homogeneity degree 1 in (a_1, a_2). So it suffices to consider the above inequality subject to $a_1^2 + a_2^2 = 1$. Then the question becomes one of showing that the maximum of $a_1 b_1 + a_2 b_2$ as a function of (a_1, a_2) subject to $a_1^2 + a_2^2 = 1$ is smaller than or equal to $\sqrt{b_1^2 + b_2^2}$. We can certainly investigate this by using the method of Lagrange multipliers; in fact, we can do this for the general case of n using a similar reduction.

(iii). But we can also investigate (1.2) directly by noting a geometric interpretation of both sides: the constraint $a_1^2 + a_2^2 = 1$ means that (a_1, a_2) is varying on the unit circle centered at the origin. Both sides of (1.2) are also of homogeneity degree 1 in (b_1, b_2), so it does not lose any information if we assume $b_1^2 + b_2^2 = 1$. Then $|a_1 b_1 + a_2 b_2|$ is the distance from the origin to the line through (a_1, a_2) with (b_1, b_2) as unit normal--geometrically, the family of straightlines $\{(a_1, a_2) : a_1 b_1 + a_2 b_2 = c\}$ for varying c all have (b_1, b_2) as their unit normal, so its maximum value is 1 and is attained when this line is tangent to the unit circle, in which case $(a_1, a_2) = \pm(b_1, b_2)$.

Illustration of the extreme values of $a_1 b_1 + a_2 b_2$ on $a_1^2 + a_2^2 = 1$ is provided at this Desmos page[1].

Note that this interpretation can also be adapted to the general n cases. If we vary two variables, say, a_i, a_j, for some $i < j$, and keep all other variables fixed, we may even keep $a_i^2 + a_j^2 = r^2$ as a constant so that the right hand side does not change. But the dependence of the left hand side on a_i, a_j is only through $a_i b_i + a_j b_j$. So the question is reduced to the two variables case, and we conclude that the left hand side attains its maximum possible value when the two factors on the right hand side are constrained to be a constant when $(a_i, a_j) = \pm(b_i, b_j)$ for any pairs $i \neq j$. Then a little further argument shows that they all have the take the same signs. Thus the left hand side attains its maximum possible value when the two factors on the right hand side are constrained to be a constant when $(a_1, \cdots, a_n) = \pm(b_1, \ldots, b_n)$.

General n cases. We set out to discuss the small n cases first, but find quickly that the ideas for solving the small n cases can easily adapt to the general n cases.

Example 1.10 Identify the maximum of $\left(\sum_{i=1}^n a_i\right)\left(\sum_{i=1}^n a_i^{-1}\right)$ under the constraint $0 < a \leq a_i \leq A$.

This can be formulated as finding the maximum of the continuously differentiable function

$$f(\mathbf{a}) := \left(\sum_{i=1}^n a_i\right)\left(\sum_{i=1}^n a_i^{-1}\right)$$

on the bounded closed set $B := \{\mathbf{a} : a \leq a_i \leq A, 1 \leq i \leq n\}$.

The maximum of f is guaranteed to exist, and the solution method is standard, but some additional analysis is needed in carrying out the solution.

First, let's see whether f has any interior critical point.

$$f_{a_k}(\mathbf{a}) = \left(\sum_{i=1}^n a_i^{-1}\right) - \left(\sum_{i=1}^n a_i\right) a_k^{-2},$$

so a critical point of f inside B must satisfy $a_1 = \cdots = a_n$. But at such a point $f(\mathbf{a}) = n^2$, and we know by the Cauchy-Schwarz inequality that for any $\mathbf{a} \in B$,

$$f(\mathbf{a}) \geq \left(\sum_{i=1}^n \sqrt{a_i}\sqrt{a_i^{-1}}\right)^2 = n^2,$$

so such a critical point only corresponds a minimum of f, and the maximum of f must occur on a boundary point of B.

The boundary of B consists of many faces and edges, so it is not a trivial task to carry out this analysis. One can get some experience by examining

[1] `https://www.desmos.com/calculator/ppezzftthy`

the $n = 2, 3$ cases first. But our computation of the partial derivative $f_{a_k}(\mathbf{a})$ contains more information:

$$f_{a_k}(\mathbf{a}) = \left(\sum_{i \neq k} a_i^{-1}\right) - \left(\sum_{i \neq k}^{n} a_i\right) a_k^{-2},$$

so it has only one zero in $\mathbb{R}^+$, is negative before this zero, and is positive after this zero. Thus the maximum in varying $a_k \in [a, A]$ must occur at the boundary point a or A.

We now see that at a maximum point of f on B,

$$a_k = a \text{ or } A \text{ for each } k.$$

Since switching the order of the a_k's does not affect $f(\mathbf{a})$, we may assume that $\mathbf{a}$ takes the form of

$$\mathbf{a} = (a, \cdots, a, A, \cdots A),$$

with m of a's, $m = 0, 1, \cdots, n$. For such an $\mathbf{a}$,

$$f(\mathbf{a}) = [ma + (n - m)A][ma^{-1} + (n - m)A^{-1}].$$

What remains is to identify the largest possible value among these $n + 1$ candidates. It turns out that

$$m = \begin{cases} \frac{n}{2} & \text{for even } n, \\ \frac{n+1}{2} & \text{for odd } n. \end{cases}$$

We thus conclude that

$$\max_B f = \begin{cases} \left(\frac{n}{2}\right)^2 (a + A)(a^{-1} + A^{-1}) = \left(\frac{n}{2}\right)^2 (1 + \frac{A}{a})(1 + \frac{a}{A}) & \text{for even } n, \\ \left(\frac{n-1}{2}\right)^2 (\frac{n+1}{n-1}a + A)(\frac{n+1}{n-1}a^{-1} + A^{-1}) = \left(\frac{n-1}{2}\right)^2 (\frac{n+1}{n-1} + \frac{A}{a})(\frac{n+1}{n-1} + \frac{a}{A}) & \text{for odd } n. \end{cases}$$

We could have seen the dependence of the bound on $\frac{A}{a}$ by noting that $f(t\mathbf{a}) = f(\mathbf{a})$ for any $t > 0$. Although the constraint $a \leq a_i \leq A$ is not invariant under the scaling $\mathbf{a} \mapsto t\mathbf{a}$, we can replace a_i by a_i/a and A by A/a in our set up: $1 \leq a_i \leq \frac{A}{a}$.

Example 1.11 Identify the minimum of $\frac{\left|\sum_{i=1}^n a_i b_i\right|}{\left(\sum_{i=1}^n a_i^2\right)^{1/2}\left(\sum_{i=1}^n b_i^2\right)^{1/2}}$ under some constraints on $\mathbf{a}, \mathbf{b}$.

The Cauchy-Schwarz inequality tells us that the above ratio is no greater than one. It can attain 0 if no constraints are imposed. Suppose that there exist $0 < a < A, 0 \leq b < B$, and

$$a \leq a_i \leq A, b \leq b_i \leq B \quad \text{for all } i.$$

What would be the minimum of the above ratio subject to this constraint? This problem is related to #92 in Part II, Chapter 2 of Polya and Szegö's classic "Problems and Theorems in Analysis I", and is an extension of the previous example.

The ratio has homogeneity of degree 0 in $\mathbf{a}$ and $\mathbf{b}$ separately, but the constraint is not homogeneity of degree 0 in $\mathbf{a}$ or $\mathbf{b}$. On the other hand, using the homogeneity of the ratio, and dividing each a_i by a, and each b_i by b, the bounds for the new a_i's are 1 and A/a, and the bounds for the new b_i's are 1 and B/b, so we expect the result will depend on A/a and B/b. We will make this reduction from here on.

Consider

$$f(\mathbf{a},\mathbf{b}) := \frac{\sum_{i=1}^n a_i b_i}{\left(\sum_{i=1}^n a_i^2\right)^{1/2}\left(\sum_{i=1}^n b_i^2\right)^{1/2}}$$

subject to the constraint $1 \leq a_i \leq A/a$ and $1 \leq b_i \leq B/b$ for each i. We first compute the derivatives of the log of f:

$$f_{a_k}/f = \frac{b_k}{\sum_{i=1}^n a_i b_i} - \frac{a_k}{\sum_{i=1}^n a_i^2},$$

$$f_{b_k}/f = \frac{a_k}{\sum_{i=1}^n a_i b_i} - \frac{b_k}{\sum_{i=1}^n b_i^2},$$

so any critical point must satisfy $\mathbf{a} = t\mathbf{b}$ for some $t > 0$. But such a critical point would make $f = 1$, which is the maximum of f; furthermore, such a critical point may not be in the interior of the constraint set. Therefore, the minimum of f in this constraint region must occur on its boundary.

We also note that

$$f_{a_k}/f = \frac{b_k \sum_{i\neq k} a_i^2 - a_k \sum_{i\neq k} a_i b_i}{\left(\sum_{i=1}^n a_i b_i\right)\left(\sum_{i=1}^n a_i^2\right)},$$

so when varying a_k only, f_{a_k} has only one zero, is positive before the zero and negative after the zero. So, like in the previous example, at any minimum of f, we must have

$$a_k = 1 \text{ or } \frac{A}{a}; b_k = 1 \text{ or } \frac{B}{b}.$$

One complication here is that rearranging the a_k's and b_k's may cause a change in the value of f. This is handled by the following algebraic fact.

> If $0 < a_1 \leq a_2 \leq \cdots \leq a_n$ and $b_1 \geq b_2 \geq \cdots \geq b_n > 0$, then for any permutation σ of $\{1, 2, \cdots, n\}$,
>
> $$\sum_{i=1}^n a_i b_i \leq \sum_{i=1}^n a_i b_{\sigma(i)} \leq \sum_{i=1}^n a_i b_{n+1-i}.$$
>
> In other words, $\sum_{i=1}^n a_i b_{\sigma(i)}$ is the smallest when the ordering of the a_i's is the opposite of that of the $b_{\sigma(i)}$'s, and the largest when their orderings are the same.

This is seen by examining the effect of permuting any two indices. Suppose $a_i \leq a_j, b_i \geq b_j$. Then

$$(a_i b_i + a_j b_j) - (a_i b_j + a_j b_i) = (a_i - a_j)(b_i - b_j) \leq 0.$$

This algebraic fact is used in an optimization problem in queuing theory. Suppose n tasks need to be completed in a queue. The completion times for these tasks are $t_1, t_2, \cdots, t_n$. We need to find a queuing arrangement so that the total waiting time by all the tasks is minimal.

An arrangement corresponds to a permutation σ of $\{1, 2, \cdots, n\}$. The associated total waiting time is

$$\begin{aligned} & t_{\sigma(1)} + t_{\sigma(1)} + t_{\sigma(2)} + \cdots + t_{\sigma(1)} + t_{\sigma(2)} + \cdots + t_{\sigma(n)} \\ =& nt_{\sigma(1)} + (n-1)t_{\sigma(2)} + \cdots + t_{\sigma(n)}. \end{aligned}$$

According to the algebraic discussion above, the minimal of this waiting time is when the t_{σ_i}'s are arranged in ascending order.

In our situation, rearranging the a_i's or $b_{\sigma(i)}$'s would not change the denominator, so the minimum of the quotient must have them arranged in opposite ordering.

Suppose we arrange the $b_{\sigma(i)}$'s in descending order,

$$\mathbf{b} = (\frac{B}{b}, \cdots, \frac{B}{b}, 1, \cdots, 1)$$

with k of $\frac{B}{b}$'s, then we must have

$$\mathbf{a} = (1, \cdots, 1, \frac{A}{a}, \cdots, \frac{A}{a}),$$

with l of $\frac{A}{a}$'s. A further argument shows that $l = k$ (You should try to provide an argument for this statement). Thus the candidates for $\min f$ in this constraint region are from

$$\frac{k\frac{B}{b} + (n-k)\frac{A}{a}}{\sqrt{[k(\frac{B}{b})^2 + n - k][k + (n-k)(\frac{A}{a})^2]}}$$

for $k = 0, \cdots, n$. It remains to find the smallest of these $n+1$ values, or find a nearly optimal lower bound.

It turns out that, if k is allowed to be a real number (so one can use one variable calculus to find the minimum), then the minimum is

$$\frac{2}{\sqrt{\frac{AB}{ab}} + \sqrt{\frac{ab}{AB}}},$$

which is the bound given by Polya and Szegö. This analysis also shows conditions under which the inequality becomes an equality.

Hint. Setting $t = (n-k)/k$, then t ranges between 0 and ∞, and

$$\frac{k\frac{B}{b} + (n-k)\frac{A}{a}}{\sqrt{[k(\frac{B}{b})^2 + n - k][k + (n-k)(\frac{A}{a})^2]}} = \frac{\frac{B}{b} + t\frac{A}{a}}{\sqrt{[(\frac{B}{b})^2 + t][1 + t(\frac{A}{a})^2]}}$$

and it's not too hard to work out the minimum of this function of t for $0 < t < \infty$ to be the claimed result above.

Checkpoint 1.12 Identify the infimum of $\frac{u\beta + v\alpha}{\sqrt{(u\beta^2+v)(u+v\alpha^2)}}$ for (u, v) in the first quadrant. Here we assume that $\alpha, \beta > 0$.

Hint. The function has homogeneity of degree 0 in (u, v), so it suffices to examine it as a function of $t = v/u$.

Answer. $\frac{2\sqrt{\alpha\beta}}{\alpha\beta+1} = \frac{2}{\sqrt{\alpha\beta}+(\sqrt{\alpha\beta})^{-1}}$.

2 Inequalities: Basic Tools of the Trade

The central concepts in analysis involve various notions of convergence and compactness. The bulk of the actual work in analysis to handle the issues involving convergence and compactness is to assess the size of the different terms, to identify the leading order ones, and use the information effectively to draw useful conclusions. This process involves regular and judicious use of various inequalities. So a good part of our discussions will involve inequalities.

2.1 Using homogeneity and scaling

Many of the examples of the previous subsection involve concrete functions, for which we can often estimate their sizes in a more specific way; we often also need to work with general functions which are not known in detail in advance, so in estimating their sizes we would need to apply equalities or inequalities that are valid for a class of general functions. In such a situation, using homogeneity and scaling is often a helpful tool as a starting point. We already used these properties in the previous subsection, here we give a bit more discussion.Any product of powers of several quantities has a notion of degree of homogeneity for the different powers. For example, if we treat a, b as separate entities, then ab has degree 1 of homogeneity for both a and b, while a^2 has degree 2 of homogeneity for a; but if we treat a, b as components of a single vector entity (a, b), then both ab and a^2 have degree 2 of homogeneity for (a, b).

In the summations

$$\sum_{i=1}^{n} a_i b_i, \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \left(\sum_{i=1}^{n} b_i^2\right)^{1/2},$$

if we treat $\mathbf{a} = (a_1, \cdots, a_n), \mathbf{b} = (b_1, \ldots, b_n)$ as separate entities, then each sum has degree 1 of homogeneity in $\mathbf{a}$ or $\mathbf{b}$.

More formally an expression $E(\mathbf{a}, \mathbf{b}, \cdots)$ is said to have degree d of (positive) homogeneity in $\mathbf{a}$ if

$$E(\lambda\mathbf{a}, \mathbf{b}, \cdots) = \lambda^d E(\mathbf{a}, \mathbf{b}, \cdots) \quad \text{for all } \lambda > 0, \mathbf{a}, \mathbf{b}, \cdots .$$

The principle of homogeneity analysis says that if both sides of an equality or inequality have a degree of homogeneity in a certain variable, then they must have the same degree of homogeneity in that variable.

Example 2.1 Degree of homogeneity in $(a + b)^n$.

In the identity

$$(a + b)^2 = a^2 + 2ab + b^2$$

each side has degree 2 homogeneity in (a, b). While

$$(a + b)^3 = a^3 + 3a^2b^2 + b^3 \text{ can't possibly hold for all } (a, b),$$

for, $(a + b)^3 - a^3 - b^3$ has degree 3 homogeneity in (a, b), but $3a^2b^2$ degree 4 homogeneity in (a, b).

We illustrate how to use this principle to prove some inequalities, including the Cauchy-Schwarz inequality.

Example 2.2 Estimate $(x+y)^a$ in terms of x^a+y^a, where $0<a<1, x,y>0$.

Both expressions have degree a of homogeneity in (x,y), so we can exploit this. We can normalize to the situation that $x+y=1$ and find an upper and lower bound of x^a+y^a.

Since $0 \le x, y \le 1$, we have $x^a \ge x$ and $y^a \ge y$, so $x^a+y^a \ge x+y=1$. Thus $x^a+y^a \ge (x+y)^a$ holds.

To get an upper bound for x^a+y^a subject to $x+y=1$ and $x,y>0$, we can eliminate the $y=1-x$ and treat $x^a+y^a=x^a+(1-x)^a$ as a one variable function of $0 \le x \le 1$. Using calculus we easily find that it attains its maximum on [0, 1] at $x=\frac{1}{2}$, so $x^a+(1-x)^a \le 2^{1-a}$, and in general we have $x^a+y^a \le 2^{1-a}(x+y)^a$.

To summarize, for any $0<a<1, x,y>0$, there holds

$$2^{a-1}(x^a+y^a) \le (x+y)^a \le x^a+y^a.$$

Example 2.3 The Cauchy-Schwarz inequality again.

The Cauchy-Schwarz inequality states that, for any $\mathbf{a}=(a_1,\cdots,a_n), \mathbf{b}=(b_1,\ldots,b_n)$,

$$|\sum_{i=1}^n a_i b_i| \le \left(\sum_{i=1}^n a_i^2\right)^{1/2} \left(\sum_{i=1}^n b_i^2\right)^{1/2}.$$

Both sides of the inequality have degree 1 of homogeneity in $\mathbf{a}$ or $\mathbf{b}$, so it suffices to establish it for

$$\sum_{i=1}^n a_i^2 = 1 \text{ and } \sum_{i=1}^n b_i^2 = 1.$$

We apply the arithmetic -geometric inequality to each $a_i b_i$

$$|a_i b_i| \le \frac{a_i^2+b_i^2}{2},$$

then sum over i to get

$$\sum_{i=1}^n |a_i b_i| \le \frac{\sum_{i=1}^n a_i^2 + \sum_{i=1}^n b_i^2}{2} = 1,$$

which is the case of the Cauchy-Schwarz inequality when $\sum_{i=1}^n a_i^2 = \sum_{i=1}^n b_i^2 = 1$.

Checkpoint 2.4 Use the same technique to prove the Hölder's inequality. The Hölder's inequality states

$$|\sum_{i=1}^n a_i b_i| \le \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} \left(\sum_{i=1}^n |b_i|^{p'}\right)^{1/p'},$$

for any vectors $\mathbf{a}=(a_1,\cdots,a_n), \mathbf{b}=(b_1,\ldots,b_n)$, where $p>1$ and p' satisfy

$\frac{1}{p} + \frac{1}{p'} = 1$.

Hint. First use convexity of $x \mapsto x^p$ (or calculus) to establish, for $a, b \geq 0$,

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

You should also consult Professor Ocone's discussion on these inequalities.

Likewise, the integrals $\int_a^b f(x)\,dx$, $\int_a^b f(x)^2\,dx$, $\left(\int_a^b f(x)^2\,dx\right)^{1/2}$ and $\left(\int_a^b f'(x)^2\,dx\right)^{1/2}$ have degrees of homogeneity in f equal to $1, 2, 1, 1$ respectively. The integrals $\int_a^b f(x)g(x)\,dx$ and $\left(\int_a^b f(x)^2\,dx\right)^{1/2}\left(\int_a^b g(x)^2\,dx\right)^{1/2}$ have degree 1 of homogeneity in f, g individually, and degree 2 of homogeneity in (f, g).

Using similar techniques, one can easily prove the integral version of the Cauchy-Schwarz inequality,

$$\Big|\int_a^b f(x)g(x)\,dx\Big| \leq \left(\int_a^b f(x)^2\,dx\right)^{1/2}\left(\int_a^b g(x)^2\,dx\right)^{1/2}$$

and Hölder's inequality,

$$\Big|\int_a^b f(x)g(x)\,dx\Big| \leq \left(\int_a^b \Big|f(x)\Big|^p\,dx\right)^{1/p}\left(\int_a^b \Big|g(x)\Big|^{p'}\,dx\right)^{1/p'},$$

where $p, p' \geq 1$, $\frac{1}{p} + \frac{1}{p'} = 1$.

Checkpoint 2.5 Prove $\int_a^b |f(x)|\,dx \leq (b-a)^{1-\frac{1}{p}}\left(\int_a^b \Big|f(x)\Big|^p\,dx\right)^{1/p}$**.** Here $p > 1$ and $f(x)$ is any function such that $|f(x)|^p$ is (Riemann) integrable on $[a, b]$. You may assume that $|f(x)|$ is also (Riemann) integrable on $[a, b]$, although this can be proved based on the given assumption.

The above discussion on homogeneity is about the scaling in the dependent variable. The scaling of the independent variable also plays a role. For instance, if D is a (nice) domain in $\mathbb{R}^n$, then we will see that $\|f\|_{L^p(D)} := \left(\int_D \Big|f(x)\Big|^p\,dx\right)^{1/p}$ scales like $(length)^{\frac{n}{p}}$, and $\|\nabla f\|_{L^p(D)} := \left(\int_D \Big|\nabla f(x)\Big|^p\,dx\right)^{1/p}$ scales like $(length)^{\frac{n-p}{p}}$. These are based on the heuristic reasoning that the volume of a domain in $\mathbb{R}^n$ scales like $(length)^n$ and taking a derivative scales like $(length)^{-1}$.

More formally, we make the change of (independent) variable $x = Ly$, and take an $f(x)$ which is *supported in* D---let's assume that D has the property that $0 \in D$ and if $x \in D$ then $x/L \in D$ for any $L \geq 1$. Then for any $L \geq 1$, $f(Ly)$ is also a function supported in D, and

$$\left(\int_D \Big|f(Ly)\Big|^p\,dy\right)^{1/p} = L^{-\frac{n}{p}}\left(\int_D \Big|f(x)\Big|^p\,dx\right)^{1/p},$$

$$\left(\int_D \Big|\nabla_y[f(Ly)]\Big|^p\,dy\right)^{1/p} = L^{\frac{p-n}{p}}\left(\int_D \Big|\nabla f(x)\Big|^p\,dx\right)^{1/p}.$$

Suppose that we believe that for a given $p \geq 1$ there is a constant $C > 0$ depending on D and some exponent q such that for all $C^1(D)$ functions $f(x)$ in D with compact support in D there holds

$$\left(\int_D \Big|f(x)\Big|^q\,dx\right)^{1/q} \leq C\left(\int_D \Big|\nabla f(x)\Big|^p\,dx\right)^{1/p}. \tag{2.1}$$

Then the same inequality should also hold with $f(Lx)$ replacing $f(x)$ for any $L \geq 1$ (let's assume that D has the property that $0 \in D$ and if $x \in D$ then $x/L \in D$ for any $L \geq 1$). This leads to

$$L^{-\frac{n}{q}} \left(\int_D \left| f(x) \right|^q dx \right)^{1/q} \leq C L^{\frac{p-n}{p}} \left(\int_D \left| \nabla f(x) \right|^p dx \right)^{1/p}.$$

In order for this inequality to hold for all $L \geq 1$, we conclude that a necessary condition is

$$-\frac{n}{q} \leq \frac{p-n}{p} \quad \text{equivalently} \quad \frac{1}{q} \geq \frac{1}{p} - \frac{1}{n}.$$

Of course this reasoning only gives a necessary condition for (2.1) to hold; it does not give an idea whether (2.1) holds. (2.1) does hold and is called the **Sobolev inequality**, but its proof would require other ideas.

Another issue related to scaling in (2.1) is the dependence of C on D. It turns out that if $\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$, then C depends only on the dimension n and is independent of D---let's denote it as S_n, while if $\frac{1}{q} > \frac{1}{p} - \frac{1}{n}$, then C has the form of $S_n |D|^{\frac{1}{q} - \frac{1}{p} + \frac{1}{n}}$, where $|D|$ is the volume of D.

Proof of the above statement. Let q_n be determined by $\frac{1}{q_n} = \frac{1}{p} - \frac{1}{n}$. Then, for q such that $\frac{1}{q} > \frac{1}{p} - \frac{1}{n}$, we know $q < q_n$. We apply Hölder's inequality to estimate $\left(\int_D \left| f(x) \right|^q dx \right)^{1/q}$ in terms of $\left(\int_D \left| f(x) \right|^{q_n} dx \right)^{1/q_n}$, then apply (2.1) with q_n to estimate the latter by $\left(\int_D \left| \nabla f(x) \right|^p dx \right)^{1/p}$.

Here is the actual implementation.

$$\begin{aligned} \left(\int_D \left| f(x) \right|^q dx \right)^{1/q} &\leq |D|^{\frac{1}{q} - \frac{1}{q_n}} \left(\int_D \left| f(x) \right|^{q_n} dx \right)^{1/q_n} \\ &\leq |D|^{\frac{1}{q} - \frac{1}{q_n}} S_n \left(\int_D \left| \nabla f(x) \right|^p dx \right)^{1/p}. \end{aligned}$$

■

Remark 2.6

An easier and more natural formulation of (2.1) which exhibits the dependence of C on the domain D is to make both sides "non-dimensionalized" *with respect to the x variable. As it stands, the two integrals in (2.1) depend on the choice of a unit for x and scale differently on this unit, so in order to make (2.1) valid in any choice of scale, the constant C has to keep track of the dependence on the scale.*

On the other hand,

$$\left(|D|^{-1} \int_D \left| f(x) \right|^q dx \right)^{1/q} \quad \text{and} \quad \left(|D|^{-1+\frac{p}{n}} \int_D \left| \nabla f(x) \right|^p dx \right)^{1/p}$$

do not depend on the choice of a scale on x, and (2.1) can be reformulated as

$$\left(|D|^{-1} \int_D \left| f(x) \right|^q dx \right)^{1/q} \leq C' \left(|D|^{-1+\frac{p}{n}} \int_D \left| \nabla f(x) \right|^p dx \right)^{1/p}$$

for some constant $C' > 0$ which is independent of the choice of a scale for D. It turns out that in the case here C' does not depend on other geometric features of D either and depends only on the dimension n.

Another often used interpretation of this formulation and the change of variable $x = Ly$ is to treat this change of variable as transforming the quantities for $x \in D$ to ones for y in some appropriatly scaled, often normalized, *region; say, if D is a ball of radius L, then $x = Ly$ would make y to lie in a unit ball, and* once a certain equality or inequality can be established on a unit ball, it can be used to establish an appropriate one on any sized ball by this scaling.

Checkpoint 2.5 can be proved in this fashion by first proving it on a unit interval and then transforming the general case to the unit interval case.

2.2 Making good use of the FTC and integration-by-parts

The Fundamental Theorem of Calculus (FTC) and integration-by-parts are essential tools in analyzing integrals involving a function and its derivatives. We illustrate their basic usages through some simple examples.

Example 2.7 (August 2014 WQ).

Let $f(x)$ be a continuously differentiable real-valued function over $\mathbb{R}$ with $f(0) = 0$. Suppose that $|f^{'}(x)| \leq |f(x)|$ for all $x \in \mathbb{R}$.

(a). Show that $f(x) = 0$ for all x in a neighborhood $(-\epsilon, \epsilon)$ for some $\epsilon > 0$.

(b). Show that $f(x) = 0$ for all $x \in \mathbb{R}$.

Solution. We should focus on how to make effective use of the assumption $|f^{'}(x)| \leq |f(x)|$ for all $x \in \mathbb{R}$. Whenever $f(x) \neq 0$, the assumption is equivalent to $|(\ln|f(x)|)'| \leq 1$, which can be readily used. How to deal with the possibility that $|f(x)|$ transits between 0 and positive? What role does the assumption $f(0) = 0$ play?

Since the assumption is easier to yield information when $f(x) \neq 0$, we will make an argument by contradiction and assume that there exists some x^* near 0 such that $f(x^*) \neq 0$ and explore the behavior of $f(x)$ in a neighborhood of x^* where $f(x) \neq 0$.

First proof. Let I be the largest interval containing x^* such that $f(x) \neq 0$ for all $x \in I$. Then for any $x \in I$, $|(\ln|f(x)|)'| \leq 1$, and, since $f(0) = 0$, I must have at least one finite end point x_e at which $f(x_e) = 0$. (Can you write down a more formal proof of this statement?)

Using the FTC, for any $x, y \in I$,

$$\begin{aligned}\Big|\ln\frac{|f(x)|}{|f(y)|}\Big| =& \Big|\ln|f(x)| - \ln|f(y)|\Big| \\ =& \Big|\int_y^x (\ln|f(s)|)'\,ds\Big| \\ \leq& |y - x|,\end{aligned}$$

from which we obtain

$$e^{-|y-x|} \leq \frac{|f(x)|}{|f(y)|} \leq e^{|y-x|}, \quad \text{for any } x, y \in I.$$

Now if we let $x \to x_e$, then $\frac{|f(x)|}{|f(y)|} \to 0$, but this would contradict the above inequality.

Second proof. Since the integral of $f'(x)$ would give the change in $f(x)$, using $f(0) = 0$, we have, for $x > 0$,

$$|f(x)| = |f(x) - f(0)| = |\int_0^x f'(s)\,ds| \leq \int_0^x |f(s)|\,ds.$$

But the derivative of $F(x) := \int_0^x |f(s)|\,ds$ is $|f(x)|$, so this has led to

$$F'(x) \leq F(x) \text{ for } x > 0.$$

This can be recognized to lead to

$$\left(e^{-x}F(x)\right)' = e^{-x}\left(F'(x) - F(x)\right) \leq 0 \text{ for } x > 0.$$

We therefore conclude that

$$e^{-x}F(x) \leq e^0 F(0) = 0 \text{ for } x > 0.$$

But this leads to

$$F(x) = \int_0^x |f(s)|\,ds \leq 0 \text{ for } x > 0,$$

which forces $f(x) = 0$ for all $x > 0$. The case for $x < 0$ can be proved in a similar fashion.

Example 2.8 Prove $|f(b) - f(a)| \leq \left(\int_a^b |f'(x)|^p\,dx\right)^{\frac{1}{p}} (b-a)^{1-\frac{1}{p}}$.

Here, $f(x)$ is any continuous function on $[a, b]$ with piecewise continuous derivative.

This is proved by applying the FTC and and Hölder's inequality.

$$|f(b){-}f(a)| = |\int_a^b f'(x)\,dx| \leq \int_a^b |f'(x)|\,dx \leq \left(\int_a^b |f'(x)|^p\,dx\right)^{\frac{1}{p}} (b{-}a)^{1-\frac{1}{p}}.$$

Remark 2.9

As a consequence of the above simple inequality, we see that, if we define

$$X := \{f \in C[a, b] : f(a) = 0, f'(x) \text{ piecewise continuous on } [a, b]\},$$

then for any $p, q \geq 1, f \in X$,

$$||f||_{L^q[a,b]} \leq (b-a)^{1-\frac{1}{p}+\frac{1}{q}} ||f'||_{L^p[a,b]}. \tag{2.2}$$

Another simple consequence of the above simple inequality is that, for any $p \geq 1, M > 0$, *the set of functions* f *in*

$$Y := \{f \in C[a, b] : f'(x) \text{ piecewise continuous on } [a, b]\}$$

with $||f'||_{L^p[a,b]} \leq M$ *is **equicontinuous**. In particular **Arzelà–Ascoli Theorem** implies that the set of functions* f *in* X *with* $||f'||_{L^p[a,b]} \leq M$ *is **pre-compact** in* $C[a, b]$, *namely, its closure in* $C[a, b]$ *is **compact** there. Put another way, for any sequence* $\{f_k\}$ *in this set, it has a subsequence* $\{f_{k_l}\}$, *and a limiting function* $f_\infty \in C[a, b]$ *such that* $\{f_{k_l}\} \to f_\infty$ *in* $C[a, b]$ *as* $l \to \infty$.

Here is the definition of *an equicontinuous family* of functions.

Definition 2.10

A family $\mathcal{F}$ of real-valued, continuous functions on a metric space X is said to be *equicontinuous* at $x_0 \in X$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\left| f(x_0) - f(y) \right| < \epsilon \quad \text{for all } f \in \mathcal{F}, \text{ whenever } d(x_0, y) < \delta.$$

The family $\mathcal{F}$ is said to be equicontinuous if it is equicontinuous at all $x \in X$.

Here is the statement of the Arzelà–Ascoli Theorem.

Theorem 2.11

Let $\mathcal{S}$ be compact. A closed subset $\mathcal{F}$ of $C(\mathcal{S})$ is compact if and only if

(i). $\sup_{\mathcal{F}} \|f\|_\infty < \infty$, *($\mathcal{F}$ is uniformly bounded), and,*

(ii). *$\mathcal{F}$ is equicontinuous.*

Professor Ocone's notes have a sketch of proof, while Professor Mirek's notes have a formulation of the notion of equicontinuity and the Arzelà–Ascoli Theorem which does not require $\mathcal{S}$ to be a metric space.

Example 2.12 Why do we care about inequalities such as (2.2)?

Inequalities about a general class of functions, perhaps subject to some side conditions such as some boundary conditions, are essential for investigating the behavior of solutions of differential equations and for their constructions.

We illustrate here a simple application of (2.2) in studying the behavior of solutions of

$$u \in C^2[a, b]; u''(x) = h(x), a < x < b; u(a) = u(b) = 0.$$

Here $h \in C[a, b]$ is treated as a given function.

The essence of (2.2) is that the L^q norm of a function is controlled by some L^p norm of the derivative of this function, provided that this function equals 0 at one end (in fact, it suffices that it equals 0 somewhere in the interval).

In our problem, the second derivative $u''(x)$ of $u(x)$ on $[a, b]$ is controlled in terms of $h(x)$. Can we control $u(x)$ and $u'(x)$ on $[a, b]$ in terms of $h \in C[a, b]$?

We would like to first apply (2.2) to $u'(x)$ on $[a, b]$, but it requires $u'(x) = 0$ somewhere in $[a, b]$. This is guaranteed by Rolle's Theorem based on the boundary conditions $u(a) = u(b) = 0$. Thus

$$\max_{[a,b]} |u'| \leq \|u''\|_{L^1[a,b]} = \|h\|_{L^1[a,b]}.$$

Next we apply (2.2) to $u(x)$ on $[a, b]$ to conclude that

$$\max_{[a,b]} |u| \leq \|u'\|_{L^1[a,b]} \leq (b - a) \max_{[a,b]} |u'| \leq (b - a)\|h\|_{L^1[a,b]}.$$

Thus $\|h\|_{L^1[a,b]}$ alone bounds $\max_{[a,b]} |u|$ and $\max_{[a,b]} |u'|$.

Question: Can one draw the same conclusions if only one boundary condition of the problem above is kept? Namely, do the same conclusions hold for solutions to

$$u \in C^2[a,b]; u''(x) = h(x), a < x < b; u(a) = 0.$$

Remark 2.13

A variant of the above inequality is the following. Let $\bar{f}$ denote the mean of f over $[a,b]$:

$$\bar{f} = (b-a)^{-1} \int_a^b f(x)\,dx.$$

Then we have, for any $x \in [a,b]$,

$$|f(x) - \bar{f}| \le \left(\int_a^b |f'(x)|^p \,dx \right)^{\frac{1}{p}} (b-a)^{1-\frac{1}{p}}. \tag{2.3}$$

One way to prove this is to use the theorem of the mean to find some $c \in [a,b]$ such that $f(c) = \bar{f}$, then apply the previous inequality on the interval between x and c.

An integral version of this inequality takes the form of

$$\|f - \bar{f}\|_{L^q[a,b]} \le (b-a)^{1-\frac{1}{p}+\frac{1}{q}} ||f'||_{L^p[a,b]}, \tag{2.4}$$

where $p, q \ge 1$.

Example 2.14 An interpolation inequality using Taylor's formula.

Suppose that $f \in C^2[a,b]$. Then for any $0 < \epsilon \le (b-a)/4$, there holds

$$\max_{[a,b]} |f'(x)| \le \epsilon \max_{[a,b]} |f''(x)| + \epsilon^{-1} \max_{[a,b]} |f(x)|.$$

One typical application of such an inequality is to obtain estimates on the derivatives of a solution to a differential equation in terms of estimates on $\max_{[a,b]} |f(x)|$. For example, if $u(x)$ solves

$$u''(x) = (10 \sin x) u'(x) - 2u(x) \text{ on } [a,b],$$

then we can use the above inequality with $\epsilon = \frac{1}{20}$ and the differential equation to obtain

$$\begin{aligned} \max_{[a,b]} |u'(x)| \le & \epsilon \max_{[a,b]} |u''(x)| + \epsilon^{-1} \max_{[a,b]} |u(x)| \\ \le & \frac{1}{2} \max_{[a,b]} |u'(x)| + \left(\frac{1}{10} + 20 \right) \max_{[a,b]} |u(x)|, \end{aligned}$$

from which we obtain

$$\max_{[a,b]} |u'(x)| \le 41 \max_{[a,b]} |u(x)|.$$

This can then be used in the differential equation to obtain estimate on $\max_{[a,b]} |u''(x)|$ in terms of $\max_{[a,b]} |u(x)|$.

Solution. This is proved by using Taylor's formula at any $x \in [a,b]$ by choosing $h = \pm 2\epsilon$ such that the interval with x and $x+h$ as ends lies in

$[a, b]$. The Taylor's formula gives

$$f(x+h) - f(x) = f'(x)h + \frac{h^2}{2} f''(c) \quad \text{for some } c \text{ between } x \text{ and } x+h.$$

It follows that

$$\begin{aligned} |f'(x)| \leq & \frac{|f(x+h) - f(x)|}{|h|} + \frac{|h| \max_{[a,b]} |f''(y)|}{2} \\ & \leq \frac{2 \max_{[a,b]} |f(y)|}{|h|} + \frac{|h| \max_{[a,b]} |f''(y)|}{2}, \end{aligned}$$

which readily gives us the desired conclusion.

Note that if $[a, b]$ is infinitely long and the natural replacements of the quantities on the right hand side, $\sup_{(a,b)} |f(x)|, \sup_{(a,b)} |f''(x)| < \infty$, then we can vary $\epsilon > 0$ arbitrarily, and by taking the $\epsilon > 0$ which minimizes the right hand side, and find the following interpolation inequality

$$\sup_{(a,b)} |f'(x)| \leq 2 \sqrt{\sup_{(a,b)} |f''(x)|} \sqrt{\sup_{(a,b)} |f''(x)|}.$$

Often we are interested in finding the optimal constant which makes (2.4) true for all f, namely, the smallest constant $C_{p,q} > 0$ such that

$$\|f - \bar{f}\|_{L^q[a,b]} \leq C_{p,q} ||f'||_{L^p[a,b]} \quad \forall f \in Y.$$

We will give an easier to work with formulation in the case of $p = q = 2$. Then the question is equivalent to identifying

$$C_N := \inf \left\{ \frac{\int_a^b |f'(x)|^2 \, dx}{\int_a^b |f(x)|^2 \, dx} : f \in Y, \bar{f} = 0 \right\}. \tag{2.5}$$

According to (2.4) for the case of $p = q = 2$, $C_N \geq (b-a)^{-2}$. Our goal is to identify C_N more explicitly, and more ambitiously, to identify those f which attain this optimal constant C_N. This is part of **Calculus of Variations**. This part is conceptually more advanced, and is typically not considered as part of the undergraduate mathematics curriculum, although the main ideas and most calculations are fairly elementary.

It turns out that

$$C_N = \left(\frac{2\pi}{(b-a)} \right)^2 \tag{2.6}$$

and any function f which attains C_N must be of the form of

$$f(x) = A \sin \left(\frac{2\pi(x-a)}{(b-a)} \right) + B \cos \left(\frac{2\pi(x-a)}{(b-a)} \right)$$

for some constants A, B.

A proof of this statement requires new ideas. The elementary methods used earlier would not work---considering that it's not clear how π would enter an elementary inequality argument and how these *minimizers* arise. We will only touch on briefly some ideas in proving such an inequality.

We will first use Fourier series expansion to sketch a proof a weaker statement, requiring the functions f to satisfy the *additional requirement* that $f(a) = f(b)$. Integration-by-parts plays a crucial role behind the scene in this proof. We first review some facts on Fourier series to be used.

Fact 2.15 Some relevant facts of Fourier series.

For ease of notation, we set $b-a=2l, a=0$.

1. *The set of functions $\{1, \cos\left(\frac{n\pi x}{l}\right), \sin\left(\frac{n\pi x}{l}\right) : n \in \mathbb{N}\}$ are mutually orthogonal to each other on $[0,2l]$ in the sense that*

$$\int_0^{2l} X(x)Y(x)\,dx = 0 \quad \text{for any distinct } X(x), Y(x) \text{ in this family.}$$

2. *Any integrable function $g(x)$ on $[0,2l]$ has a Fourier expansion*

$$g(x) \sim a_0 + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{n\pi x}{l}\right) + b_n \sin\left(\frac{n\pi x}{l}\right)\right],$$

where a_n, b_n are obtained formally by multiplying both sides of the above relation by either $\cos\left(\frac{n\pi x}{l}\right)$ or $\sin\left(\frac{n\pi x}{l}\right)$ and integrating over $[0,2l]$. Using the above orthogonal relation, we have

$$a_0 = \frac{1}{2l}\int_0^{2l} g(x)\,dx,$$

$$a_n = \frac{1}{l}\int_0^{2l} g(x)\cos\left(\frac{n\pi x}{l}\right)\,dx \text{ for } n \geq 1,$$

$$b_n = \frac{1}{l}\int_0^{2l} g(x)\sin\left(\frac{n\pi x}{l}\right)\,dx \text{ for } n \geq 1.$$

3. *The expansion relation above is an equality in the mean square sense. Namely, the partial sums*

$$S_N[g](x) := a_0 + \sum_{n=1}^{N} \left[a_n \cos\left(\frac{n\pi x}{l}\right) + b_n \sin\left(\frac{n\pi x}{l}\right)\right]$$

converge to $g(x)$ in the mean square sense on $[0,2l]$:

$$\|g - S_N[g]\|_{L^2[0,2l]} \to 0 \text{ as } N \to \infty.$$

4. *Furthermore, the following* Parseval equality *holds (which is a version of the Pythagorean Theorem in this context):*

$$\int_0^{2l} |g(x)|^2\,dx = 2l|a_0|^2 + l\sum_{n=1}^{\infty}\left[|a_n|^2 + |b_n|^2\right].$$

5. *Assume that $g(x)$ is piecewise C^1 on $[0,2l]$. Denote the Fourier series expansion of $g'(x)$ on $[0,2l]$ by*

$$g'(x) \sim a_0' + \sum_{n=1}^{\infty} \left[a_n' \cos\left(\frac{n\pi x}{l}\right) + b_n' \sin\left(\frac{n\pi x}{l}\right)\right].$$

Assume further that $g(x)$ is continuous *on $[0,2l]$, and $g(0)=g(2l)$. Then*

$$a_0' = 0, \text{ and } a_n' = \left(\frac{n\pi}{l}\right)b_n, b_n' = -\left(\frac{n\pi}{l}\right)a_n.$$

In other words, under our assumptions here, the Fourier series expansion of $g'(x)$ on $[0,2l]$ can be obtained by term-wise differentiation of the Fourier series expansion of $g(x)$ on $[0,2l]$.

Remark 2.16

The following properties are used in a crucial way in deriving the above properties of Fourier series.

$$
\begin{aligned}
&\int_0^{2l} \cos^2\left(\frac{n\pi x}{l}\right) dx = \int_0^{2l} \sin^2\left(\frac{n\pi x}{l}\right) dx = l \quad \textit{for } n \in \mathbb{N},\\
&\int_0^{2l} g'(x)\cos\left(\frac{n\pi x}{l}\right) dx\\
=&g(x)\cos\left(\frac{n\pi x}{l}\right)\Big|_{x=0}^{x=2l} + \left(\frac{n\pi}{l}\right)\int_0^{2l} g(x)\sin\left(\frac{n\pi x}{l}\right) dx \textit{ if } g \in C[0,2l]\\
=&\left(\frac{n\pi}{l}\right)\int_0^{2l} g(x)\sin\left(\frac{n\pi x}{l}\right) dx \textit{ if } g(0) = g(2l);\\
&\int_0^{2l} g'(x)\sin\left(\frac{n\pi x}{l}\right) dx\\
=&g(x)\sin\left(\frac{n\pi x}{l}\right)\Big|_{x=0}^{x=2l} - \left(\frac{n\pi}{l}\right)\int_0^{2l} g(x)\cos\left(\frac{n\pi x}{l}\right) dx \textit{ if } g \in C[0,2l]\\
=&-\left(\frac{n\pi}{l}\right)\int_0^{2l} g(x)\cos\left(\frac{n\pi x}{l}\right) dx.
\end{aligned}
$$

Checkpoint 2.17 Examine the relations of the Fourier series of a function and its derivative. In item 5 above the assumption that $g(x)$ is *continuous* on $[0, 2l]$ and $g(0) = g(2l)$ is essential. Evaluate the Fourier series of the following functions and their derivatives, and examine whether the relations in item 5 above hold.

1. $f(x) = x - 1$ on$[0, 2]$
2. $f(x) = \begin{cases} x & if 0 \le x \le 1, \\ x - 2 & if 1 < x \le 2 \end{cases}$

Proving a restrictive version of (2.6) using with Fourier series expansion. For any $f \in Y, \bar{f} = 0$ satisfying the additional assumption that $f(0) = f(2l)$, both f and f' has its Fourier series expansion on $[0, 2l]$:

$$
\begin{aligned}
f(x) &\sim a_0 + \sum_{n=1}^{\infty}\left[a_n \cos\left(\frac{n\pi x}{l}\right) + b_n \sin\left(\frac{n\pi x}{l}\right)\right],\\
f'(x) &\sim a'_0 + \sum_{n=1}^{\infty}\left[a'_n \cos\left(\frac{n\pi x}{l}\right) + b'_n \sin\left(\frac{n\pi x}{l}\right)\right].
\end{aligned}
$$

Note that

$$
\begin{aligned}
a_0 &= \frac{1}{2l}\int_0^{2l} f(x)\,dx = 0 && \text{using the assumption } \bar{f} = 0,\\
a'_0 &= \frac{1}{2l}\int_0^{2l} f'(x)\,dx = 0 && \text{using } f(0) = f(2l),\\
a'_n &= \left(\frac{n\pi}{l}\right) b_n && \text{also using } f(0) = f(2l),\\
b'_n &= -\left(\frac{n\pi}{l}\right) a_n.
\end{aligned}
$$

Now, according to the Parseval equality,

$$\begin{aligned}
\|f\|^2_{L^2[0,2l]} &= \int_0^{2l} |f(x)|^2\,dx = 2l|a_0|^2 + l\sum_{n=1}^{\infty}\left[|a_n|^2+|b_n|^2\right],\\
\|f'\|^2_{L^2[0,2l]} &= \int_0^{2l} |f'(x)|^2\,dx = 2l|a_0'|^2 + l\sum_{n=1}^{\infty}\left[|a_n'|^2+|b_n'|^2\right]\\
&= l\sum_{n=1}^{\infty}\left(\frac{n\pi}{l}\right)^2\left[|a_n|^2+|b_n|^2\right]\\
&\ge \left(\frac{\pi}{l}\right)^2 l\sum_{n=1}^{\infty}\left[|a_n|^2+|b_n|^2\right]\\
&= \left(\frac{\pi}{l}\right)^2 \|f\|^2_{L^2[0,2l]}.
\end{aligned}$$

Since equality holds iff $a_n = b_n = 0$ for all $n \ge 2$, namely, when $f(x) = A\cos\left(\frac{\pi x}{l}\right) + B\sin\left(\frac{\pi x}{l}\right)$, (2.6) now follows. ■

We now briefly describe the approach using calculus of variations. There are several issues to deal with.

1. Is there an f which attains this optimal constant C_N?
2. If the answer to the above is affirmative, is there a characterization for such an f?
3. Can use the information above to evaluate C_N?

We will assume that answer to item 1 above is affirmative, and briefly discuss items 2 and 3. Set

$$Q[f] := \frac{\int_a^b |f'(x)|^2\,dx}{\int_a^b |f(x)|^2\,dx} \text{ for } f \in Y, \bar{f} = 0, f \not\equiv 0.$$

First, (2.4) for the case of $p = q = 2$ shows that the infimum of $Q[f]$ is positive. Next, assume that $f \in Y, \bar{f} = 0, f \not\equiv 0$ attains the infimum, then for any $h \in Y, \bar{h} = 0$ and any t, $f + th$ is a competitor, and

$$C_N = Q[f] \le Q[f+th] \quad \forall t \text{ small so that } f + th \not\equiv 0.$$

It is routine to see that $Q[f + th]$ is a differentiable function in t, so we must have

$$\left.\frac{d}{dt}\right|_{t=0} Q[f+th] = 0.$$

Since $Q[cf] = Q[f]$ for any $c \neq 0$, we may scale f, if necessary, to make $\int_a^b |f(x)|^2\,dx = 1$. Then a routine computation shows that

$$\left.\frac{d}{dt}\right|_{t=0} Q[f+th] = 2\int_a^b \left[f'(x)h'(x) - C_N f(x)h(x)\right]\,dx.$$

Assuming that f is twice continuously differentiable, then integrating by parts on the first integral gives

$$\int_a^b f'(x)h'(x)\,dx = f'(x)h(x)\Big|_{x=a}^{x=b} - \int_a^b f''(x)h(x)\,dx.$$

Therefore we conclude that

$$f'(x)h(x)\Big|_{x=a}^{x=b} - \int_a^b [f''(x) + C_N f(x)]\, h(x)\, dx = 0 \quad \forall h \in Y, \bar{h} = 0. \tag{2.7}$$

If we take $h \in Y, \bar{h} = 0$, and further that $h(a) = h(b) = 0$, then we get

$$-\int_a^b [f''(x) + C_N f(x)]\, h(x)\, dx = 0 \quad \forall h \in Y, \bar{h} = 0, h(a) = h(b) = 0.$$

We now use the following calculus facts.

Fact 2.18

- *If $g \in C[a,b]$ such that $\int_a^b g(x)h(x)\, dx = 0$ for any $h \in C[a,b]$ with $h(a) = h(b) = 0$, then $g \equiv 0$ in $[a,b]$.*
- *If $g \in C[a,b]$ such that $\int_a^b g(x)h(x)\, dx = 0$ for any $h \in C[a,b]$ with $\bar{h} = 0$, then g equals some constant in $[a,b]$.*
- *If $g \in C[a,b]$ such that $\int_a^b g(x)\, dx = 0$ and $\int_a^b g(x)h(x)\, dx = 0$ for any $h \in C[a,b]$ with $\bar{h} = 0, h(a) = h(b) = 0$, then $g = 0$ in $[a,b]$.*

*For the first property, if $g(c) \neq 0$ for some c, construct some h*supported near*c to make $\int_a^b g(x)h(x)\, dx \neq 0$. For the second property, take any $\eta \in C[a,b]$, then take $h = \eta - \bar{\eta}$ and use*

$$\int_a^b g(x)h(x)\, dx = \int_a^b [g(x) - \bar{g}]\, \eta(x)\, dx.$$

Using Fact 2.18 and noting that, if we take $\eta \in C[a,b]$ such that $\eta(a) = \eta(b) = 0$ and $h = \eta - \bar{\eta}$ in (2.7), it takes the form of

$$\begin{aligned}
0 &= -[f'(b) - f'(a)]\bar{\eta} - \int_a^b [f''(x) + C_N f(x)]\, (\eta(x) - \bar{\eta})\, dx \\
&= -[f'(b) - f'(a)]\bar{\eta} - \int_a^b [f''(x) + C_N f(x)]\, \eta(x)\, dx + [f'(b) - f'(a)]\bar{\eta} \text{ (using } \bar{f} = 0) \\
&= -\int_a^b [f''(x) + C_N f(x)]\, \eta(x)\, dx
\end{aligned}$$

It follows that $f''(x) + C_N f(x) = 0$ on $[a,b]$.

Now (2.7) takes the form of

$$f'(x)h(x)\Big|_{x=a}^{x=b} = 0 \quad \forall h \in Y, \bar{h} = 0.$$

Since we can allow $h(a), h(b)$ arbitrary subject to $h \in Y, \bar{h} = 0$, it follows that $f'(a) = f'(b) = 0$. To summarize, any function f which attains the C_N must satisfy

$$f''(x) + C_N f(x) = 0, \quad a < x < b; \quad f'(a) = f'(b) = 0.$$

The question has been reduced to finding possible values C_N for which the above problem has a solution f which is not identically 0, and then identify the smallest such C_N. It turns out that only when $C_N = \left(\frac{n\pi}{l}\right)^2$ with $l = (b-a)/2$ and for some $n \in \mathbb{N}$, can we find some solution f; in fact, in such a case,

$$f(x) = A \cos\left(\frac{n\pi(x-a)}{l}\right)$$

for some constant A. Thus we identify C_N to be $\left(\frac{\pi}{l}\right)^2$.

Remark 2.19

Note that in the restrictive formulation of (2.6) we require the functions to satisfy $g(a) = g(b)$, and the minimizers under this condition are identified to be $A\cos\left(\frac{\pi(x-a)}{l}\right) + B\sin\left(\frac{\pi(x-a)}{l}\right)$ for some constants A, B; while in the full formulation of (2.6) we find the infimum to be the same, but the minimizers are more restrictive. In general the boundary conditions play an integral role in such problems and in problems involving differential equations.

Checkpoint 2.20 Identify $\inf Q[f]$ on a modified set of functions. Consider

$$Z_1 = \{f \in C[a,b] : f'(x) \text{ piecewise continuous on } [a,b], f(a) = 0\}$$

and

$$Z_2 = \{f \in C[a,b] : f'(x) \text{ piecewise continuous on } [a,b], f(a) = f(b) = 0\}.$$

Identify $\inf\{Q[f] : f \in Z_1\}$ and $\inf\{Q[f] : f \in Z_2\}$. Also try to identify the minimizers in each case.

3 More Applications of Integration-by-parts

We include here some more applications of integration-by-parts.

Example 3.1 Taylor's remainder formula.

For any $n+1$ times continuously differentiable function f on $[a,b]$, there exists c between a and b such that

$$f(b) = \sum_0^n \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}. \tag{3.1}$$

Solution. We will give a proof by using suitably chosen integration-by-parts repeatedly to express $f(b) - f(a)$ using higher order derivatives of f.

$$\begin{aligned}
f(b) &= f(a) + \int_a^b f'(x)\,dx \\
&= f(a) + \int_a^b f'(x)\,d(x-b) \\
&= f(a) + f'(x)(x-b)\big|_{x=a}^{x=b} - \int_a^b (x-b)f''(x)\,dx \\
&= f(a) + f'(a)(b-a) - \int_a^b f''(x)\,d\frac{(x-b)^2}{2} \\
&= f(a) + f'(a)(b-a) - f''(x)\frac{(x-b)^2}{2}\big|_{x=a}^{x=b} + \int_a^b \frac{(x-b)^2}{2}f'''(x)\,dx \\
&= f(a) + f'(a)(b-a) + f''(a)\frac{(b-a)^2}{2} + \int_a^b f'''(x)\,d\frac{(x-b)^3}{3!} \\
&= \cdots
\end{aligned}$$

$$= f(a) + f'(a)(b-a) + f''(a)\frac{(b-a)^2}{2} + \cdots + f^{(n)}(a)\frac{(b-a)^n}{n!}$$
$$+ (-1)^n \int_a^b f^{(n+1)}(x)\frac{(x-b)^n}{n!}\,dx.$$

In the second line above, $d(x-b)$ is chosen so an undesired boundary term is absent after integration by parts. This gives an integral version of the Taylor's reminder term. If we apply the theorem of the mean to the last integral, we find some $c \in (a,b)$ such that

$$\begin{aligned}
&(-1)^n \int_a^b f^{(n+1)}(x)\frac{(x-b)^n}{n!}\,dx \\
=&f^{(n+1)}(c)\int_a^b \frac{(b-x)^n}{n!}\,dx \\
=&f^{(n+1)}(c)\frac{(b-a)^{n+1}}{(n+1)!}.
\end{aligned}$$

Integration-by-parts is often used to account for cancellation in some integrals. Here are some examples.

Example 3.2 For any continuous function f on $[0,2\pi]$, $\int_0^{2\pi} f(x)\sin(nx)\,dx \to 0$ as $n \to \infty$.

Solution. When f is a constant, we see easily that $\int_0^{2\pi}\sin(nx)\,dx = 0$. This property holds not only on $[0,2\pi]$, for any $a<b$, we also have $\int_a^b \sin(nx)\,dx \to 0$ as $n\to\infty$. This is the cancellation property we referred to.

If $f \in C^1[0,2\pi]$, then we can exploit this cancellation property using integration-by-parts as follows.

$$\begin{aligned}
\int_0^{2\pi} f(x)\sin(nx)\,dx &= -\frac{1}{n}\int_0^{2\pi} f(x)\,d\cos(nx) \\
&= -\left.\frac{f(x)\cos(nx)}{n}\right|_{x=0}^{2\pi} + \frac{1}{n}\int_0^{2\pi}\cos(nx)f'(x)\,dx \\
&\to 0 \text{ as } n\to\infty.
\end{aligned}$$

Now for any continuous function f on $[0,2\pi]$, and any $\epsilon > 0$, we first find some $g \in C^1[0,2\pi]$ such that $|f(x)-g(x)| < \epsilon$ for all $x \in [0,2\pi]$. Then

$$\int_0^{2\pi} f(x)\sin(nx)\,dx = \int_0^{2\pi}[f(x)-g(x)]\sin(nx)\,dx + \int_0^{2\pi} g(x)\sin(nx)\,dx.$$

Now
$$\left|\int_0^{2\pi}[f(x)-g(x)]\sin(nx)\,dx\right| \le 2\pi\epsilon,$$
and
$$\int_0^{2\pi} g(x)\sin(nx)\,dx \to 0 \text{ as } n\to\infty,$$
so there exists some N such that
$$\left|\int_0^{2\pi} g(x)\sin(nx)\,dx\right| < \epsilon \text{ for all } n > N.$$

Thus we get

$$\left|\int_0^{2\pi} f(x)\sin(nx)\,dx\right| < (2\pi+1)\epsilon \text{ for all } n > N,$$

which shows that $\int_0^{2\pi} f(x)\sin(nx)\,dx \to 0$ as $n \to \infty$.

Note that we didn't use the continuity assumption on f directly: the proof relies on approximating f by C^1 functions, and we only need to get the approximation *in the integral sense* instead of uniformly: for a sequence of $g_k \in C^1[0, 2\pi]$

$$\int_0^{2\pi} \left|f(x) - g_k(x)\right| dx \to 0 \text{ as } k \to \infty.$$

Question: Do you know an explicit procedure to approximate a continuous function on a closed interval uniformly by a sequence of C^1 functions? How about approximating a Riemann integrable function in the integral sense by a sequence of C^1 functions?

One possibility is to work with the average $h^{-1}\int_x^{x+h} f(y)\,dy$.

Checkpoint 3.3 Prove that the improper integral $\int_0^\infty \sqrt{x}\sin(x^2)\,dx$ is convergent.

Hint. Treat the integrand $\sqrt{x}\sin(x^2)$ as $\frac{1}{2\sqrt{x}}\left(\sin(x^2)\right)'$.

It turns out that another way to handle the previous example is to exploit the periodicity of one of the factors in the integrand.

Example 3.4 For any continuous function f on $[0, 2\pi]$, $\int_0^{2\pi} f(x)|\sin(nx)|\,dx \to \frac{2}{\pi}\int_0^{2\pi} f(x)\,dx$ as $n \to \infty$.

Here we note that $|\sin(nx)|$ has period $\frac{\pi}{n}$, and

$$\int_0^{\frac{\pi}{n}} |\sin(nx)|\,dx = \frac{2}{n}.$$

When $n \to \infty$, we can partition $[0, 2\pi]$ into union of short intervals of the form $[\frac{(k-1)\pi}{n}, \frac{k\pi}{n}]$, and approximate f on this interval by its value at one end, say, $f(\frac{k\pi}{n})$, which would lead to an approximation of the integral on that interval by $\frac{2}{n}f(\frac{k\pi}{n})$. But the sum of these approximations gives us a Riemann sum for the integral of f on $[0, 2\pi]$ multiplied by a factor of $\frac{2}{\pi}$. Here are more details.

$$\begin{aligned}
&\int_0^{2\pi} f(x)|\sin(nx)|\,dx \\
=&\sum_{k=1}^{2n}\int_{\frac{(k-1)\pi}{n}}^{\frac{k\pi}{n}} f(x)|\sin(nx)|\,dx \\
=&\sum_{k=1}^{2n}\int_{\frac{(k-1)\pi}{n}}^{\frac{k\pi}{n}} [f(x) - f(\frac{k\pi}{n})]|\sin(nx)|\,dx + \sum_{k=1}^{2n}\frac{2}{n}f(\frac{k\pi}{n}).
\end{aligned}$$

On $[\frac{(k-1)\pi}{n}, \frac{k\pi}{n}]$, $[f(x) - f(\frac{k\pi}{n})]|\sin(nx)|$ is bounded by the oscillation of f on it, so when $n \to \infty$, the first term tends to 0 due to the Riemann integrability of f, while the second term tends to $\frac{2}{\pi}\int_0^{2\pi} f(x)\,dx$.

Checkpoint 3.5 Find the limit $\int_a^b f(x)s(nx)\,dx$ as $n \to \infty$ when $s(x)$ is a general periodic function. The ideas in the previous example can be generalized. Suppose that $s(x)$ has period $T > 0$. Then

$$\int_a^b f(x)s(nx)\,dx \to \left(\frac{1}{T}\int_0^T s(x)\,dx\right)\int_a^b f(x)\,dx \text{ as } n \to \infty.$$

The second mean-value theorem for the integral is often used to estimate an integral when the integrand is the product of a monotone function and another function whose integral has control. It's proof under the general condition stated below needs to work with the Riemann sum definition of integrals and use the **Abel summation-by-parts formula**. But if we assume that the monotone factor is continuously differentiable, then one can use integration-by-parts to give a simple proof.

Theorem 3.6 Second mean-value theorem for the integral.

If f, g are integrable on $[a, b]$ and g is a monotonic function on [a, b], then there exists a point $\xi \in [a, b]$such that

$$\int_a^b f(x)g(x)\,dx = g(a)\int_a^\xi f(x)\,dx + g(b)\int_\xi^b f(x)\,dx.$$

Proof. We provide a proof when g is assumed to be continuously differentiable. Set $F(x) = \int_a^x f(y)\,dy$. Then $F(x)$ is a continuously differentiable function, and

$$\begin{aligned}\int_a^b f(x)g(x)\,dx = &\int_a^b g(x)\,dF(x)\\ =&g(x)F(x)\Big|_{x=a}^{x=b} - \int_a^b F(x)g'(x)\,dx.\end{aligned}$$

Let m, M be the minimum and maximum value of $F(x)$ over $[a, b]$, respectively. Then, using that $g'(x)$ has the same sign for all $x \in [a, b]$,

$$\int_a^b F(x)g'(x)\,dx \text{ is between } m\int_a^b g'(x)\,dx \text{ and } M\int_a^b g'(x)\,dx,$$

so equals $\bar{F}\int_a^b g'(x)\,dx$ for some $m \le \bar{F} \le M$. On the other hand, there exists some $\xi \in [a, b]$ such that $\bar{F} = F(\xi)$, so we get

$$\begin{aligned}\int_a^b f(x)g(x)\,dx =&g(b)F(b) - F(\xi)\int_a^b g'(x)\,dx\\ =&g(b)F(b) - F(\xi)\,(g(b) - g(a))\\ =&g(b)\,[F(b) - F(\xi)] + g(a)F(\xi)\\ =&g(a)\int_a^\xi f(x)\,dx + g(b)\int_\xi^b f(x)\,dx.\end{aligned}$$

■

Theorem 3.7 Abel–Dirichlet test for convergence of an improper integral.

Let $x \mapsto f(x)$ and $x \mapsto g(x)$ be functions defined on an interval $[a, c)$ and integrable on every closed interval $[a, b] \subset [a, c)$, where c may be ∞. Suppose that g is monotonic. Then under one of the following pair of conditions, the integral $\int_a^c f(x)g(x)\,dx$ is convergent.

(A). The integral $\int_a^c f(x)\,dx$ is convergent, and g is bounded.

(B). The integral $\int_a^c f(x)\,dx$ is bounded, and $g(x)$ converges to 0 as $x \to c$.

Proof. One simply verifies the Cauchy integral criterion by applying the second mean-value theorem for the integral. ■

Checkpoint 3.8 Verify that $\int_0^\infty \frac{\arctan(x)\sin x}{x} dx$ is convergent.

4 Modes of Convergence

When studying convergence of a sequence of functions there are different notions of convergence. The most elementary ones are *point wise convergence* and *uniform convergence*. In applications there is often a need to deal with *convergence in the integral sense*. We will briefly discuss the relation and difference of these notions and how they are used in applications.

4.1 Definition and Motivation

Let D denote the domain of a sequence of functions f_k. We may take D to be an interval (a, b). Let $p \geq 1$.

Definition 4.1

We say that f_k converges in $L^p(D)$, if there exists a limit function $f \in L^p(D)$ such that

$$\|f_k - f\|_{L^p(D)} \to 0 \text{ as } k \to \infty.$$

It is often easier to check whether a sequence of functions converges point wise, but when it does, we often gain very little on behavior of the limit function. For example, if each f_k is Riemann integrable, and the sequence f_k converges to f point wise, it may not imply that the limit f is Riemann integrable, and even if it is, it may not imply that $\int_D |f_k(x) - f(x)|^p\,dx \to 0$ for any $p \geq 1$.

Here is an example illustrating how convergence in $L^p(D)$ arises in applications. We will briefly describe how convergence in $L^2(D)$ arises in constructing a solution to the mixed boundary-initial value problem for the heat equation

$$\begin{aligned} u_t(x,t) - u_{xx}(x,t) &= 0 & \quad 0 < x < l, t > 0, \\ u(0,t) = u(l,t) &= 0 & t > 0, \\ u(x,0) &= g(x) & 0 < x < l, \end{aligned}$$

where the initial data $g(x)$ is a given continuous function on $[0, l]$, and traditionally we would like the solution $u(x,t)$ to be twice continuously differentiable in x, once continuously differentiable in t in the domain $(0, l) \times (0, \infty)$, and continuous on $[0, l] \times [0, \infty)$.

There is an elementary procedure of looking for separable solutions of the form $X(x)T(t)$, which solves the homogeneous heat equation and the homogeneous boundary conditions. The result is that for any $n \in \mathbb{N}$,

$$\sin\left(\frac{n\pi x}{l}\right) e^{-\left(\frac{n\pi}{l}\right)^2 t}$$

is such a solution. Since we are so far dealing with linear homogeneous equations, any linear combination of solutions is still a solution, so

$$\sum_{n \in \text{a finite set}} c_n \sin\left(\frac{n\pi x}{l}\right) e^{-\left(\frac{n\pi}{l}\right)^2 t}$$

also satisfies the same equations. What remains is whether one can choose the c_n's so that this solution at $t = 0$ gives rise to the prescribed initial data $g(x)$. For that purpose, first we need to form an infinite sum and demand that

$$\sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi x}{l}\right) = g(x) \quad \text{on } (0, l) \text{ in an appropriate sense.}$$

But we also need to make sense of the infinite series as a continuously differentiable solution. This is a version of the Fourier series expansion. It turns out that we must choose c_n such that

$$c_n = \frac{2}{l} \int_0^l g(x) \sin\left(\frac{n\pi x}{l}\right) \, dx.$$

We will only focus on the issue of in what sense

$$u(x,t) := \sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi x}{l}\right) e^{-\left(\frac{n\pi}{l}\right)^2 t} \to g(x) \text{ for } 0 < x < l, \text{ as } t \to 0+.$$

It turns out that it would take considerable effort to prove that $u(x,t)$ is continuous for $(x,t) \in [0,l] \times [0,\infty)$, that $u(y,t) \to g(x)$ as $y \to x \in (0,l), t \to 0+$, and one also needs to impose $g(0) = g(l) = 0$ to prove that the convergence is uniform over $x \in (0,l)$. On the other hand, it is fairly easy, and natural, to prove that $u(x,t) \to g(x)$ in the mean square sense, namely, $\|u(x,t) - g(x)\|_{L^2(0,l)} \to 0$ as $t \to 0+$.

This is seen by using the Parseval equality: for any $t, t' \geq 0$,

$$||u(\cdot,t) - u(\cdot,t')||^2_{L^2[0,l]} = \frac{2}{l} \sum_{n=1}^{\infty} c_n^2 \left| e^{-\left(\frac{n\pi}{l}\right)^2 t} - e^{-\left(\frac{n\pi}{l}\right)^2 t'} \right|^2;$$

$$\frac{2}{l} \sum_{n=1}^{\infty} c_n^2 = \|g\|^2_{L^2(0,l)}.$$

Now for any given $\epsilon > 0$, we can find N such that $\sum_{n=N+1}^{\infty} c_n^2 < \frac{\epsilon}{8}$, which then leads to

$$\sum_{n=N+1}^{\infty} c_n^2 \left| e^{-\left(\frac{n\pi}{l}\right)^2 t} - e^{-\left(\frac{n\pi}{l}\right)^2 t'} \right|^2 \leq 4 \sum_{n=N+1}^{\infty} c_n^2 \leq \frac{\epsilon}{2} \text{ for } t, t' \geq 0;$$

On the other hand, there exists $\delta > 0$ such that, when $|t - t'| < \delta$,

$$\sum_{n=1}^{N} c_n^2 \left| e^{-\left(\frac{n\pi}{l}\right)^2 t} - e^{-\left(\frac{n\pi}{l}\right)^2 t'} \right|^2 \leq \frac{\epsilon}{2},$$

which proves that $u(\cdot,t)$ is (uniformly) continuous in $L^2(0,l)$, including at $t = 0$.

4.2 An Integral Convergence Theorem and a Generalization

When (a, b) is a bounded interval (or when D has bounded volume), uniform convergence implies convergence in $L^p(D)$. But when (a, b) is not a bounded interval (or when D has infinite volume), uniform convergence does not necessarily imply convergence in $L^p(D)$; additional control is needed to obtain convergence in $L^p(D)$. This is in reference to the following commonly used integral convergence theorem.

Theorem 4.2 Integral Convergence Theorem.

Let f_k be a sequence of continuous functions on the bounded interval $[a, b]$. Assume f_k converges uniformly to f on $[a, b]$. Then

$$\lim_{n\to\infty} \int_a^b f_k(s)\, ds = \int_a^b f(s)\, ds.$$

Example 4.3 A Uniformly Convergent Sequence May Fail to Converge in $L^p(D)$.

consider $f_k(x) = \frac{k}{k^2+x^2}$ on $[1, \infty)$. Since $0 < f_k(x) \le \frac{k}{k^2+1}$ for all $k \ge 1, x \in [1, \infty)$, it is clear that $f_k(x) \to 0$ uniformly on $[1, \infty)$. However, using the change of variables $x = ky$, we see that

$$\int_{[1,\infty)} |f_k(x) - 0|\, dx = \int_{[1,\infty)} \frac{k}{k^2+x^2}\, dx = \int_{[\frac{1}{k},\infty)} \frac{1}{1+y^2}\, dy \to \frac{\pi}{2},$$

so $f_k(x)$ does not converge to 0 in $L^1([1, \infty))$.

It is clear that the failure of convergence in $L^1[1, \infty)$ is due to lack of uniform control on $\int_N^\infty f_k\, dx$: fix any $\epsilon > 0$, for each k, we can find some N_k such that $\int_N^\infty f_k\, dx < \epsilon$; but there is no such N that would work for all k. In other words, there exists some $\epsilon_0 > 0$, such that for any k, there is some $N_k \to \infty$, $\int_{N_k}^\infty f_k\, dx \ge \epsilon_0$.

In many situations we can identify one or a finite number of locations near which a sequence of functions f_k may fail to converge uniformly, but away from such points, f_k converges uniformly; or f_k converges uniformly over a set D, but D may fail to have finite volume, the additional conditions needed to guarantee the convergence of f_k in $L^p(D)$ often come in the form of some control of f_k near these identified points. We will give below a generalization of the above integral convergence theorem; a more general theorem (Lebesgue's Dominated Convergence Theorem) will be one of the main results in the first year graduate analysis course.

Theorem 4.4 A Generalized Integral Convergence Theorem.

In the following we allow $b = \infty$. Let f_k be a sequence of continuous functions on the interval (a, b). Assume that for any $c, a < c < b$, f_k converges uniformly to f on (a, c). Assume further that there exists a function $g(x)$, integrable on (a, c) for any $c, a < c < b$, such that (i). the integral (perhaps improper) $\int_a^b g(x)\, dx$ is convergent, (ii). $|f_k(x)| \le g(x)$ for any $x \in (a, b)$. Then the integral $\int_a^b f(s)\, ds$ is convergent and

$$\lim_{n\to\infty} \int_a^b f_k(s)\, ds = \int_a^b f(s)\, ds.$$

In fact, the stronger statement

$$\lim_{n\to\infty}\int_a^b |f_k(s)-f(s)|\,ds = 0$$

holds.

Proof. First we show that the integrals $\int_a^b f_k(s)\,ds, \int_a^b f(s)\,ds$ are convergent. It suffices to show that for any $\epsilon > 0$, there exists some $c, a < c < b$, such that for any $c', c < c' < b$,

$$|\int_c^{c'} f_k(s)\,ds|, |\int_c^{c'} f(s)\,ds| \leq \epsilon.$$

This can be done by first identifying some $c, a < c < b$, such that for any $c', c < c' < b$, $|\int_c^{c'} g(s)\,ds| \leq \epsilon/2$. It follows that $|\int_c^{c'} f_k(s)\,ds| \leq \epsilon/2$ for all k. For any $c', c < c' < b$, apply the regular integral convergence theorem above to f_k on (c, c'), we get $|\int_c^{c'} f(s)\,ds| \leq \epsilon$.

Next we show the limiting integral property along a similar line. It is clear that point wise we have $|f(x)| \leq g(x)$. We use the same set up of the above paragraph, then

$$\begin{aligned}
&\int_a^b |f_k(s)-f(s)|\,ds \\
\leq &\int_a^c |f_k(s)-f(s)|\,ds + \int_c^b |f_k(s)-f(s)|\,ds \\
\leq &\int_a^c |f_k(s)-f(s)|\,ds + 2\int_c^b g(s)\,ds \\
\leq &\int_a^c |f_k(s)-f(s)|\,ds + \epsilon.
\end{aligned}$$

Now we apply the regular integral convergence theorem above to f_k on (a, c) to get some N such that for all $k \geq N$, $\int_a^c |f_k(s)-f(s)|\,ds < \epsilon$. This concludes our proof. ■

Example 4.5 Differentiation under the integral $\int_0^1 |x-y|^\alpha \rho(y)\,dy$.

Here $0 < \alpha < 1$ and ρ is Riemann integrable on $[0,1]$. We will show that $u(x) := \int_0^1 |x-y|^\alpha \rho(y)\,dy$ is differentiable in $x \in (0,1)$ and

$$u'(x) = \int_0^1 \alpha(x-y)|x-y|^{\alpha-2}\rho(y)\,dy.$$

Namely we can differentiate under the integral sign here.

The main issue is that $(x-y)|x-y|^{\alpha-2} \to \infty$ as $y \to x$, so we are dealing with an improper integral here. Furthermore, if $h_k \to 0$, then, in examining the difference quotient

$$\frac{u(x+h_k)-u(x)}{h_k} = \int_0^1 \frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}\rho(y)\,dy,$$

if we fix any $y \neq x$, then

$$\frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}\rho(y) \to \alpha(x-y)|x-y|^{\alpha-2}\rho(y) \text{ as } k \to \infty,$$

but this convergence is not uniform over the set $y \neq x$. However, for any $\delta > 0$, the convergence is uniform over $(0,1) \setminus (x-\delta, x+\delta)$. So the main issue is the behavior for y near x. For simplicity, we split the integral as $\int_0^x + \int_x^1$ and only work out some details for the integral $\int_0^x$.

It turns out that it is not even easy to apply our generalized integral convergence theorem, as the difference quotient $\frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}$ has its absolute value equal to $|h_k|^{\alpha-1}$ at $y = x, x+h_k$, which $\to \infty$ as $k \to \infty$, so it is not easy to find a function $g(x)$ satisfying the conditions in the generalized integral convergence theorem. However, we can adapt the ideas in the proof of that theorem to handle the situation here. We will break the integral $\int_0^x$ into three pieces $\int_0^{x-\delta} + \int_{x-\delta}^{x-|h_k|} + \int_{x-|h_k|}^x$.

Note that when $y < x - |h_k|$, we have, by the theorem of the mean,

$$\Big|\frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}\Big| = \alpha|x+\theta_k h_k - y|^{\alpha-1} \leq \alpha\Big|x - |h_k| - y\Big|^{\alpha-1},$$

for some $0 < \theta_k < 1$, so

$$\Big|\int_{x-\delta}^{x-|h_k|} \frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}\rho(y)\,dy\Big| \leq \int_{x-\delta}^{x-|h_k|} \Big|x-|h_k|-y\Big|^{\alpha-1}|\rho(y)|\,dy.$$

Since there exists some $M > 0$ such that $|\rho(y)| \leq M$, and the improper integral

$$\int_{x-\delta}^{x-|h_k|} \alpha\Big|x - |h_k| - y\Big|^{\alpha-1}\,dy \leq \delta^\alpha,$$

it follows that, for any given $\epsilon > 0$, we can take $\delta > 0$ small enough to make

$$\Big|\int_{x-\delta}^{x-|h_k|} \frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}\rho(y)\,dy\Big| \leq \epsilon.$$

On the other hand,

$$\begin{aligned}&\Big|\int_{x-|h_k|}^{x} \frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}\rho(y)\,dy\Big| \\ \leq&\frac{M}{|h_k|}\int_{x-|h_k|}^{x} (|x+h_k-y|^\alpha + |x-y|^\alpha)\,dy \leq \frac{M}{|h_k|}(2|h_k|)^{\alpha+1},\end{aligned}$$

so for $h_k \to 0$, we see that for sufficiently large k,

$$\Big|\int_{x-\delta}^{x} \frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}\rho(y)\,dy\Big| \leq 2\epsilon.$$

It remains to examine the limit

$$\int_0^{x-\delta} \frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}\rho(y)\,dy,$$

but we can deal with this limit using the regular integral convergence theorem. To put things together, we have

$$\begin{aligned}&\Big|\int_0^x \frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k}\rho(y)\,dy - \int_0^x \alpha(x-y)|x-y|^{\alpha-2}\rho(y)\,dy\Big| \\ \leq&\int_0^{x-\delta}\Big|\frac{|x+h_k-y|^\alpha - |x-y|^\alpha}{h_k} - \alpha(x-y)|x-y|^{\alpha-2}\Big||\rho(y)|\,dy\end{aligned}$$

$$+\int_{x-\delta}^{x}\Big|\frac{|x+h_k-y|^\alpha-|x-y|^\alpha}{h_k}-\alpha(x-y)|x-y|^{\alpha-2}\Big||\rho(y)|\,dy$$
$$\le\int_0^{x-\delta}\Big|\frac{|x+h_k-y|^\alpha-|x-y|^\alpha}{h_k}-\alpha(x-y)|x-y|^{\alpha-2}\Big||\rho(y)|\,dy+4\epsilon.$$

Finally, we can find some N such that for $k\ge N$,

$$\int_0^{x-\delta}\Big|\frac{|x+h_k-y|^\alpha-|x-y|^\alpha}{h_k}-\alpha(x-y)|x-y|^{\alpha-2}\Big||\rho(y)|\,dy\le\epsilon,,$$

which leads to

$$\Big|\int_0^x\frac{|x+h_k-y|^\alpha-|x-y|^\alpha}{h_k}\rho(y)\,dy-\int_0^x\alpha(x-y)|x-y|^{\alpha-2}\rho(y)\,dy\Big|\le 5\epsilon.$$

This complete an "$\epsilon-\delta$" type argument for showing

$$\lim_{k\to\infty}\Big|\int_0^x\frac{|x+h_k-y|^\alpha-|x-y|^\alpha}{h_k}\rho(y)\,dy-\int_0^x\alpha(x-y)|x-y|^{\alpha-2}\rho(y)\,dy\Big|=0.$$

5 Completeness and Compactness

Completeness is so essential in analysis that its role can't be emphasized enough. The main advantage of Lebesgue's integration theory is that the space of Lebesgue integrable functions is complete in the integral norm, in contrast to the space of Riemann integrable functions. Compactness is another essential concept in analysis. Its role is to reduce a problem involving infinitely many possibilities to a finite number of possibilities.

5.1 Statements of the Bolzano-Weierstrass Theorem and Heine-Borel Theorem

In Modern literature, the Bolzano-Weierstrass Theorem refers to the following theorem.

Theorem 5.1 Bolzano-Weierstrass Theorem.

Every bounded sequence in $\mathbb{R}^n$ contains a convergent subsequence.

Definition 5.2 Compactness and sequential compactness.

A set K in a topological space X (you may take X to be a metric space) is called **compact**, if any cover of K by open sets of X has a *finite subcover*.

A set S in X is called **sequentially compact**, if any sequence $(x_k)\subset S$ has a convergent subsequence with limit in S.

A set S in X is called **sequentially pre-compact**, if any sequence $(x_k)\subset S$ has a convergent subsequence with limit in X.

Using Theorem Theorem 5.1, it is easy to prove

Theorem 5.3 Bolzano-Weierstrass Theorem (Alternate Version).

A subset K in $\mathbb{R}^n$ is sequentially pre-compact iff it is bounded.
A subset K in $\mathbb{R}^n$ is sequentially compact iff it is bounded and closed.

Heine-Borel Theorem refers to the following theorem.

Theorem 5.4 Heine-Borel Theorem.

A subset K in $\mathbb{R}^n$ is compact iff it is bounded and closed.

It follows from Bolzano-Weierstrass Theorem and Heine-Borel Theorem that subset in $\mathbb{R}^n$ is compact iff it is sequentially compact.

Historically Bolzano formulated and proved a result, which he used to prove the intermediate-value-property of a continuous function on a closed interval, and is equivalent to the following form of completeness of the set of real numbers.

Theorem 5.5 Bolzano's Lemma.

Any nonempty set of $\mathbb{R}$ bounded above has a least upper bound.

Heine's version of the Heine-Borel Theorem was about the finite subcovering property of a closed interval, and he used it to prove that a continuous function defined on a bounded closed interval is uniformly continuous.

Checkpoint 5.6 Reconstruct proofs of the uniform continuity of a continuous function on a bounded closed interval, one using the Bolzano-Weierstrass Theorem, and another using the Heine-Borel Theorem. Then extend your proof to a continuous function on a compact metric space.

Question: How do the Bolzano-Weierstrass Theorem and Heine-Borel Theorem generalize in more general contexts?

Theorem 5.7 Heine-Borel Theorem in a metric space.

A set K in a metric space X is compact iff it is sequentially compact; both are equivalent to the condition that K be complete and totally bounded.

Note that a bounded set in $\mathbb{R}^n$ is totally bounded, but a bounded set in a general metric space may not be totally bounded.

Checkpoint 5.8 Prove that the unit ball in l^p is not totally bounded. Can you prove the same statement for the unit ball in a general infinite-dimensional normed space?

Professor Ocone's notes contain a sketch of proof of Heine-Borel Theorem in a metric space.

Question 5.9 In Professor Ocone's proof of "compactness $\Longrightarrow$ sequential compactness", where did he use the metric of the space?

Professor Ocone's proof of "compactness $\Longrightarrow$ sequential compactness" implies that any infinite set S of points in a compact set must contain an *accumulation* point of S , defined as a point such that within any neighborhood there are infinitely many points of S. When applied to a sequence of points in S, and the sequence has infinitely many points, the argument produces an accumulation point of the sequence. How does this give rise to a convergent subsequence?

Remark 5.10

When the topology can't be given via a metric (such a space is called non-metrizable*), compactness and sequential compactness may not be equivalent to each other. In most applications in analysis we use sequential compactness.*

In Elementary analysis, the following two characterizations of the continuity of a map $f : X \subset \mathbb{R}^n \mapsto Y$ at some $x_0 \in X$ are equivalent

(a) *For any open neighborhood V of $f(x_0)$ in Y, $f^{-1}(V)$ is an open neighborhood of x_0 in X.*

(b) *For any sequence $(x_k) \to x_0$ in X, $f(x_k) \to f(x_0)$ as $k \to \infty$.*

This equivalence still holds in a metric space. In a general topological space, (a) still implies (b), but (b) may not imply (a). For example, under the weak topology of l^1 (we will motivate the notion of weak topology through an example in the next subsection, but will not have time to discuss it), the function $f(x) = ||x||_{l^1} := \sum_{n=1}^{\infty} |x(n)|$ for $x = (x(n)) \in l^1$ satisfies (b), but not (a). The verification of (b) uses the so called Schur's Lemma, which implies that a weak convergent sequence in l^1 is also norm convergent in l^1.

Baire's category theorem is an important property of a complete metric space, and has many applications. Professor Mirek's notes provide some discussions and examples. The January 2011 Qualifying Exam[1] has a problem which can be solved using the Baire's category theorem. We will not have time to discuss this theorem.

5.2 Some Examples

We will first discuss an example illustrating the difference between the roles of the notions of completeness and compactness.

Example 5.11 (August, 2012 WQ).

Suppose that (X, d) is a complete metric space with a finite diameter: i.e. there exists $D < \infty$ such that

$$d(x, y) \leq D \text{ for all } x, y \in X.$$

Is it true that every continuous function f on X is bounded? Prove this assertion or give a counterexample.

Solution. If we review a proof for the property that every continuous function f on X is bounded, a key ingredient is that X be compact. A complete metric space with a finite diameter is a bounded closed set, but we only know that a bounded closed set of a finite dimensional Euclidean space is compact.

Due to the Bolzano-Weierstrass Theorem in any finite dimensional Euclidean space, any failure of the above property can only occur in an infinite dimensional setting. If we take X to be the subset of functions in $C[a, b]$ with supremum norm ≤ 2, then it is a complete metric space with finite

[1] `https://math.rutgers.edu/docman-lister/math-main/academics/graduate/qualifying-exam/1267-wqw2011/file`

diameter. We examine whether it is possible to have a certain continuous function defined on X, when evaluated at a sequence u_k in X, becomes unbounded. The key is that this sequence u_k should not have any convergent subsequence. If we take a unit step function v defined on $[a, b]$, then it is not continuous, and can't be approximated uniformly over $[a, b]$ by continuous functions in X, but can be approximated by continuous functions in X in the integral sense, thus if we define

$$f[u] = (\|u - v\|_{L^1[a,b]})^{-1} \text{ for } u \in X,$$

then it is well defined, and there exists a sequence u_k in X bounded in the sup norm, such that $\|u - v\|_{L^1[a,b]} \to 0$, which implies that $f[u_k] \to \infty$.

Question: Can you construct a similarly behaving example in some other spaces such as l^p?

Checkpoint 5.12 (August 2012 WQ). Let X and Y be locally compact metric spaces, and let $f : X \mapsto Y$ be a continuous mapping which is bijective. Show that

$$f \text{ is a homeomorphism} \Leftrightarrow f^{-1}(K) \text{ is compact for all compact } K \subset Y.$$

Note: a metric space is locally compact if and only if every point has an open neighborhood with compact closure.

Checkpoint 5.13 Any two norms on a finite dimensional vector space are equivalent. Let X be any finite dimensional vector space over $\mathbb{R}$, and let $\mathbf{v}_1, \cdots, \mathbf{v}_n$ be a basis. For any $\mathbf{x} \in X$, let $(\mathbf{x}(1), \cdots, \mathbf{x}(n))$ be the coordinates of $\mathbf{x}$ in this basis:

$$\mathbf{x} = \sum_{k=1}^{n} \mathbf{x}(k)\mathbf{v}_k.$$

Then there exits constants $c_2 > c_1 > 0$ such that

$$c_2 \left(\sum_{k=1}^{n} |\mathbf{x}(k)|^2 \right)^{1/2} \geq ||\mathbf{x}|| \geq c_1 \left(\sum_{k=1}^{n} |\mathbf{x}(k)|^2 \right)^{1/2} \quad \forall \mathbf{x} \in X.$$

Hint. The first inequality follows from triangle inequality. For the second inequality, use homogeneity and examine effect of a sequence $\mathbf{x}_m$ such that

$$\left(\sum_{k=1}^{n} |\mathbf{x}_m(k)|^2 \right)^{1/2} = 1, \text{ but } ||\mathbf{x}_m|| \to 0.$$

We will next briefly describe an example illustrating the need for working with the space of functions complete in the L^2 norm. It concerns the **variational method** of constructing a solution of the boundary value problem

$$u''(x) = f(x), 0 < x < l, \tag{5.1}$$
$$u(0) = u(l) = 0. \tag{5.2}$$

For an initial value problem, namely replacing the two boundary conditions by the initial conditions of prescribing $u(0), u'(0)$, we have a standard procedure for constructing a solution. Professor Ocone's notes provide a review of using a fixed point argument to construct a solution. That procedure also relies on the completeness of the space of continuous functions in the sup norm.

It turns out that a solution of the boundary value problem (5.1) (5.2) can be identified as a *minimizer* of the *functional*

$$E[u] := \int_0^l \left(\frac{1}{2}|u'(x)|^2 + f(x)u(x)\right) dx$$

defined on the space X of continuously differentiable functions on $[0, l]$ such that $u(0) = u(l) = 0$. Namely, if u is a solution of (5.1) (5.2), then

$$E[u] \le E[v] \text{ for any } v \in X; \tag{5.3}$$

and conversely, if $u \in X$ satisfies (5.3), and u is twice continuously differentiable in $(0, l)$ (this can actually be proved based on the previous property), then it is a solution of (5.1)(5.2).

The first claim can be seen by setting $w = v - u$, and noting that

$$\begin{aligned}
&E[v] = E[u + w] \\
&= \int_0^l \left(\frac{1}{2}|u'(x)|^2 + \frac{1}{2}|w'(x)|^2 + u'(x)w'(x) + f(x)u(x) + f(x)w(x)\right) dx \\
&= E[u] + \frac{1}{2}\int_0^l |w'(x)|^2\, dx + \int_0^l (u'(x)w'(x) + f(x)w(x))\, dx,
\end{aligned}$$

lastly integrating-by-parts on the last integral

$$\int_0^l (u'(x)w'(x) + f(x)w(x))\, dx = u'(x)w(x)|_{x=0}^{x=l} + \int_0^l (-u''(x) + f(x))\, w(x)\, dx = 0,$$

using $w(0) = w(l) = 0$ and $-u''(x) + f(x) = 0$.

The second claim can be seen by noting that a consequence of (5.3) is

$$0 = \frac{d}{dt}\Big|_{t=0} E[u + tw] = \int_0^l (u'(x)w'(x) + f(x)w(x))\, dx,$$

and that integrating-by-parts on this integral gives

$$\begin{aligned}
0 &= \int_0^l (u'(x)w'(x) + f(x)w(x))\, dx \\
&= u'(x)w(x)|_{x=0}^{x=l} + \int_0^l (-u''(x) + f(x))\, w(x)\, dx, \text{ for any } w \in X,
\end{aligned}$$

which then implies that

$$-u''(x) + f(x) = 0$$

for $x \in (0, l)$.

We now discuss the role played by completeness in proving the existence of some u satisfying (5.3). First we argue that

$$m := \inf_X E[v] \text{ is finite.}$$

This is seen by applying (2.2) and Hölder's inequality as follows.

$$\begin{aligned}
E[v] &\ge \int_0^l \frac{1}{2}|v'(x)|^2\, dx - ||f||_{L^2(0,l)}||v||_{L^2(0,l)} \\
&\ge \frac{1}{2}||v'||^2_{L^2(0,l)} - l||f||_{L^2(0,l)}||v'||_{L^2(0,l)} \\
&\ge \frac{1}{2}||v'||^2_{L^2(0,l)} - \frac{1}{4}||v'||^2_{L^2(0,l)} - l^2||f||^2_{L^2(0,l)}
\end{aligned}$$

$$\geq \frac{1}{4}||v'||^2_{L^2(0,l)} - -l^2||f||^2_{L^2(0,l)}$$

which shows that m is finite. Next we take a minimizing sequence for $E[v]$, namely, $u_j(x) \in X$ such that $E[u_j] \to m$. The above argument shows that $||u_j'||_{L^2(0,l)}$ is bounded. Furthermore, noting

$$E[u_j] + E[u_k] = 2E[\frac{u_j + u_k}{2}] + \frac{1}{4}||u_j' - u_k'||^2_{L^2(0,l)} \geq 2m + \frac{1}{4}||u_j' - u_k'||^2_{L^2(0,l)}$$

and $E[u_j], E[u_k] \to m$, we conclude that $||u_j' - u_k'||^2_{L^2(0,l)} \to 0$ as $j, k \to \infty$. Namely, the sequence $\{u_j\}$ is a Cauchy sequence in the norm $||u_j'||_{L^2(0,l)}$!

The issue is that X *is not complete* under this norm! The way to resolve this issue is to define the **completion** Y of X under this norm, extend $E[v]$ to Y, and rework the previous steps in the space Y. One final step is to show that a minimizer u in Y is still twice continuously differentiable on $(0, l)$, therefore is a solution to (5.1)(5.2).

Next we use an example to illustrate the notion of compactness in an infinite dimensional space. More examples of applications of compactness will come in the examples of next section.

Example 5.14 A compactness criterion for a subset of l^p.

Recall that, for $1 \leq p < \infty$, the space l^p consists of the set of infinite sequences $\{x = (x(1), x(2), \cdots)\}$ such that $\|x\|_p := \left(\sum_{k=1}^{\infty} |x(k)|^p\right)^{1/p} < \infty$. l^p is a complete normed space with $\|x\|_p$ as its norm.

It is easy to see that the closed unit ball

$$B := \{x \in l^p : \|x\|_p \leq 1\}$$

of l^p is not compact, as the sequence

$$\{e_k = (0, \cdots, 0, 1, 0, \cdots) : \text{ with the only 1 on the } k^{th} \text{ slot}\}$$

is in B, but can't have any convergent subsequence.

We now give a compactness criterion for a subset of l^p.

> A subset K of l^p is compact iff
>
> 1. K is a bounded and closed subset of l^p;
> 2. For any $\epsilon > 0$, there exists some N such that
>
> $$\left(\sum_{k=N+1}^{\infty} |x(k)|^p\right)^{1/p} < \epsilon \text{ for all } x \in K.$$

Item 2 above says that the sum of the "tail part" of $x \in K$ can be made uniformly small (so we only need to focus on the first N components of $x \in K$).

We will use the compactness criterion for a set in a metric space, which is reviewed in Professor Ocone's notes. More specifically,

> a set K in a metric space is compact iff K is complete and totally bounded. "Totally bounded" means that for every $\epsilon > 0$, K can be covered by a finite number of open balls of radius less than or equal to ϵ.

We will only sketch a proof that if K satisfies our criterion, then it is totally bounded. Consider

$$K_N := \{(x(1), x(2), \cdots, x(N)) : x \in K\}.$$

It is a bounded and closed subset of $\mathbb{R}^N$ under our assumptions, so is compact in $\mathbb{R}^N$. Thus there exists a finite number L of points $\{x_1, \cdots, x_L\}$ in this subset such that any $x \in K$, $y = (x(1), x(2), \cdots, x(N)) \in K_N$ is in some ball $B_\epsilon(x_l)$ of radius ϵ centered at some x_l, $1 \le l \le L$. Now for any $x \in K$, we apply the Minkowski inequality to imply

$$\begin{aligned}\|x - x_l\|_p \le & \left(\sum_{k=1}^{N} |x(k) - x_l(k)|^p\right)^{1/p} + \left(\sum_{k=N+1}^{\infty} |x(k) - x_l(k)|^p\right)^{1/p} \\ \le & \epsilon + \left(\sum_{k=N+1}^{\infty} |x(k)|^p\right)^{1/p} + \left(\sum_{k=N+1}^{\infty} |x_l(k)|^p\right)^{1/p} \\ \le & 3\epsilon,\end{aligned}$$

which shows that K is totally bounded.

Remark 5.15

While the above example indicates that a key characteristic of a compact set K in l^p is that the sum of the "tail part" of $x \in K$ can be made uniformly small (so we only need to focus on the first N components of $x \in K$), the analogue of this property of a compact set $\mathcal{F}$ in $C(\mathcal{S})$, where $\mathcal{S}$ is a compact metric space, is that $\mathcal{F}$ be equicontinuous.

Roughly speaking, this property controls the size of oscillation *of the functions in $\mathcal{F}$ in a common neighborhood, so when*

$$\begin{aligned}&|f(y) - f(x_0)| < \epsilon \text{ for all } y, d(y, x_0) < \delta, f \in \mathcal{F} \text{ holds, and} \\ &\{f(x_0) : f \in \mathcal{F}\} \text{ is covered by a finite number of balls of radius } \epsilon \\ &\text{centered at } f_1, \cdots, f_N, \text{ then} \\ &\forall y, d(y, x_0) < \delta, f \in \mathcal{F}, |f(y) - f_k| < 2\epsilon \text{ for some } k.\end{aligned}$$

This is how we reduce the uniform convergence of a sequence of functions in $\mathcal{F}$ to the convergence of this sequence at a finite number of points.

Checkpoint 5.16 The set of points $\{x \in l^2 : \sum_{l=1}^{\infty} l|x(l)|^2 \le 1\}$ is compact in l^2.

Remark 5.17

*The above example provides a good illustration for the notion of **(sequential) weak convergence**. Let $\{x_k\} \subset l^p, p > 1$, be a sequence such that for each fixed index l, the sequence of the l^{th} components $\{x_k(l)\}_{k=1}^{\infty} \to y(l)$ for some $y(l)$, and there exists some $M > 0$ such that*

$$\|x_k\|_p \le M \text{ for all } k.$$

It is easy to see that it does not necessarily follow that $\|x_k - y\|_p \to 0$ as $k \to \infty$, where $y = (y(1), y(2), \cdots)$. However, for any $z \in l^q$, where q is the *conjugate exponent of* $p: p^{-1}+q^{-1}=1$, *we will show that*

$$\langle x_k, z\rangle := \sum_{l=1}^{\infty} x_k(l)z(l) \to \sum_{l=1}^{\infty} y(l)z(l) = \langle y, z\rangle.$$

This may be considered as an earliest example motivating the notion of (sequential) weak convergence. Our claim amounts a generalization of the Bolzano-Weierstrass compactness criterion to this infinite dimensional context: Any sequence bounded in $l^p, p>1$ *has a subsequence which converges weakly in the above sense.*

The claim is seen by using z *to control the sum of the "tail" part: for any* $\epsilon > 0$, *there exists some* N *such that*

$$\left(\sum_{l=N+1}^{\infty} z(l)^q\right)^{1/q} < \epsilon.$$

Then

$$\begin{aligned}
&|\langle x_k - y, z\rangle| \\
\leq& |\sum_{l=1}^{N}[x_k(l)-y(l)]z(l)| + |\sum_{l=N+1}^{\infty}[x_k(l)-y(l)]z(l)| \\
\leq& |\sum_{l=1}^{N}[x_k(l)-y(l)]z(l)| + \left(\sum_{l=N+1}^{\infty}|x_k(l)-y(l)|^p\right)^{1/p}\left(\sum_{l=N+1}^{\infty}|z(l)|^q\right)^{1/q} \\
\leq& |\sum_{l=1}^{N}[x_k(l)-y(l)]z(l)| + \left[\left(\sum_{l=N+1}^{\infty}|x_k(l)|^p\right)^{1/p} + \left(\sum_{l=N+1}^{\infty}|y(l)|^p\right)^{1/p}\right]\left(\sum_{l=N+1}^{\infty}|z(l)|^q\right)^{1/q} \\
\leq& |\sum_{l=1}^{N}[x_k(l)-y(l)]z(l)| + [M+M]\,\epsilon.
\end{aligned}$$

Here we have used

$$\left(\sum_{l=m}^{L}|y(l)|^p\right)^{1/p} = \left(\lim_{k\to\infty}\sum_{l=m}^{L}|x_k(l)|^p\right)^{1/p} \leq M,$$

for any finite $m < L$ *to imply*

$$\left(\sum_{l=m}^{\infty}|y(l)|^p\right)^{1/p} \leq M.$$

(You will see a generalization of this in the first year graduate analysis course in the form of Fatou's Lemma.) Finally, our assumption implies that

$$|\sum_{l=1}^{N}[x_k(l)-y(l)]z(l)| \to 0 \text{ as } k \to \infty,$$

so we have

$$|\sum_{l=1}^{N}[x_k(l)-y(l)]z(l)| < \epsilon \text{ for all sufficiently large } k.$$

This concludes the proof that $\langle x_k, z\rangle \to \langle y, z\rangle$ *as* $k \to \infty$.

Remark 5.18

For any $p, 1 \le p < \infty$, and for any $\mathbf{y} \in l^{p'}$, where p' is the conjugate exponent of $p : p^{-1} + p'^{-1} = 1$, then by Hölder's inequality

$$f_{\mathbf{y}}(\mathbf{x}) := \langle \mathbf{y}, \mathbf{x} \rangle = \sum_{n=1}^{\infty} x(n) y(n)$$

defines a continuous linear functional *on l^p, namely, a continuous function $f : l^p \mapsto \mathbb{R}$ such that*

$$f(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2) = a_1 f(\mathbf{x}_1) + a_2 f(\mathbf{x}_2)$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in l^p, a_1, a_2 \in \mathbb{R}$. It turns out that any such a continuous linear functional is equal to $f_{\mathbf{y}}$ for some $\mathbf{y} \in l^{p'}$. For this reason, the space $l^{p'}$ is called the dual space *of l^p.*

The weak topology on l^p is defined through the union of finite intersections of the special open sets of the form

$$\{\mathbf{x} : a < f_{\mathbf{y}}(\mathbf{x} - \mathbf{x}_0) < b\}$$

for some $\mathbf{x}_0 \in l^p, \mathbf{y} \in l^{p'}, a < b \in \mathbb{R}$. Fix any $\mathbf{x}_0 \in l^p$, by varying $\mathbf{y} \in l^{p'}, a < 0 < b \in \mathbb{R}$ and taking finite intersections of sets of the above form, we obtain a neighborhood base *for $\mathbf{x}_0 \in l^p$ in the weak topology of l^p. Thus any neighborhood of $\mathbf{x}_0 \in l^p$ in the weak topology of l^p contains an open set of the form*

$$\{\mathbf{x} : a_k < f_{\mathbf{y}_k}(\mathbf{x} - \mathbf{x}_0) < b_k, 1 \le k \le N\}$$

for some N and $a_k < 0 < b_k, 1 \le k \le N$. This only imposes a finite number of linear constraints (more properly, affine constraints).

This definition is related to the definition of product topology *of a family of topological spaces.*

Since l^p is infinite dimensional, for any finite number N of $\mathbf{y}_k \in l^{p'}$, there exists some non-zero *$\mathbf{v} \in l^p$ such that $f_{\mathbf{y}_k}(\mathbf{v}) = 0$ for all $k = 1, \cdots, N$. Thus for any $c \in \mathbb{R}$, vectors of the form $\mathbf{x}_0 + c\mathbf{v}$ will be in the open neighborhood $\{\mathbf{x} : a_k < f_{\mathbf{y}_k}(\mathbf{x} - \mathbf{x}_0) < b_k, 1 \le k \le N\}$ of $\mathbf{x}_0$, so no neighborhood in the weak topology of l^p in bounded in the norm of l^p.*

6 Convexity and Some Applications

Convexity plays an important role in many extremal problems and inequalities. Professor Ocone's notes illustrate some basic applications (mostly in dimension one), including in proving the Hölder and Minkowski inequalities. We will add some additional discussion here.

6.1 Convex sets vs Convex Functions

The notion of a convex set in $\mathbb{R}^n$ is more general than that of a convex function.

Definition 6.1

A set C in $\mathbb{R}^n$ is called convex, if for any $\mathbf{a}, \mathbf{b} \in C$ and any $t \in \mathbb{R}, 0 \le t \le 1$, we have $(1 - t)\mathbf{a} + t\mathbf{b} \in C$.

Geometrically, the set $\{(1-t)\mathbf{a}+t\mathbf{b} : 0 \le t \le 1\}$ is the line segment in $\mathbb{R}^n$ with $\mathbf{a}, \mathbf{b}$ as its ends. A convex set needs not have any interior point.

Definition 6.2

A real-valued function f defined on a convex set C is called convex, if for any $\mathbf{a}, \mathbf{b} \in C$ and any $t \in \mathbb{R}, 0 \le t \le 1$, we have

$$f((1-t)\mathbf{a}+t\mathbf{b}) \le (1-t)f(\mathbf{a}) + tf(\mathbf{b}).$$

f is called strictly convex if we have the strict inequality

$$f((1-t)\mathbf{a}+t\mathbf{b}) < (1-t)f(\mathbf{a}) + tf(\mathbf{b}) \text{ for any } 0 < t < 1.$$

f is called concave if $-f$ is convex. Equivalently, the defining inequality above is reversed for a concave function.

Geometrically, if we construct a line in $\mathbb{R}^n \times \mathbb{R} \supset C \times \mathbb{R}$ through the points $(\mathbf{a}, f(\mathbf{a})), (\mathbf{b}, f(\mathbf{b}))$, then it has parametric equation

$$\mathbf{x} = ((1-t)\mathbf{a}+t\mathbf{b}, (1-t)f(\mathbf{a}) + tf(\mathbf{b})),$$

so $(1-t)f(\mathbf{a})+tf(\mathbf{b})$ is the "height of the line above the point" $(1-t)\mathbf{a}+t\mathbf{b}$. When f is convex, $f((1-t)\mathbf{a}+t\mathbf{b})$ stays below the height at $(1-t)\mathbf{a}+t\mathbf{b}$ of the above line segment for $0 \le t \le 1$. Here[1] is a Desmos page illustrating this geometric property.

f is a convex function iff the set $\{(\mathbf{x}, y) : \mathbf{x} \in C, y \ge f(\mathbf{x})\}$ in $\mathbb{R}^n \times \mathbb{R}$, called the **epigraph** of f, is convex. Another characterization of a convex function is that for every real number c, the **sub level set** of f defined by $\{\mathbf{x} : f(\mathbf{x}) \le c\}$ is a convex set.

Because of this relation, properties of convex functions can often be studied as properties of convex functions. We will later discuss briefly the notion of **supporting hyperplane** of a convex set and that of the graph of a convex function.

6.2 Some properties and applications of univariate convex functions

The illustration above in the previous subsection also includes a sketch of the argument that the slope of the secant lines on a convex function of a single variable is an increasing function. Geometrically it seems clear that if f is a convex function of a single variable, $a < a < c$, then $f(c)$ stays above the secant line through $(a, f(a), (b, f(b))$. This can be derived from the above property of secant lines: for $c > b$,

$$\frac{f(c)-f(a)}{c-a} \ge \frac{f(b)-f(a)}{b-a} \Leftrightarrow f(c) \ge f(a) + \frac{f(b)-f(a)}{b-a}(c-a).$$

These inequalities also hold when $c < a$.

The continuity of a convex function of one variable at an interior point is proved using these bounds by linear functions. Say, a is an interior point. Then there exist b, c in the domain such that $b > a > c$, and the above property of secant lines implies that for any $x, b > x > a$,

$$f(a) + \frac{f(c)-f(a)}{c-a}(x-a) \le f(x) \le f(a) + \frac{f(b)-f(a)}{b-a}(x-a).$$

[1] https://www.desmos.com/calculator/mbdd9ymzug

Then the sandwich theorem implies that $f(x) \to f(a)$ as $x \to a+$. The direction when $x \to a-$ is done in a similar way.

Question 6.3 How can we extend this argument to higher dimensions? The function is certainly continuous when constrained along any one-dimensional lines, but there are infinitely many lines through any given point. Here we will see some ideas of compactness at play.

Solution. We will assume that the origin is in the interior of the domain of f; in fact, we will assume the domain of f includes the unit ball centered at the origin, and describe ideas to prove the continuity of f at the origin.

The key idea is that *an appropriate choice of n points can be used to form its* ***convex hull****,*

$$\{t_1\mathbf{a}_1 + \cdots + t_n\mathbf{a}_n : t_i \geq 0, t_1 + \cdots + t_n = 1\},$$

whose projection on the unit sphere $\mathbb{S}^{n-1}$ covers an non-empty open set. As a result

$$f(t_1\mathbf{a}_1 + \cdots + t_n\mathbf{a}_n) \leq t_1 f(\mathbf{a}_1) + \cdots + t_n f(\mathbf{a}_n) \leq M,$$

where M is chosen so that $f(\mathbf{a}_i) \leq M, i = 1, \cdots, n$. In fact, if we take $\mathbf{a}_i = \mathbf{e}_i$, the standard basis vector in $\mathbb{R}^n$, then

$$\begin{aligned} Q_+ :=& \{t_1\mathbf{a}_1 + \cdots + t_n\mathbf{a}_n : t_i \geq 0, t_1 + \cdots + t_n = 1\} \\ =& \{(t_1, \cdots, t_n) : t_i \geq 0, t_1 + \cdots + t_n = 1\}. \end{aligned}$$

The next idea is to bound $f(\mathbf{x})$ from above for those $\mathbf{x}$ which lie on a segment between the origin and a point in Q_+. In fact, for any $\mathbf{x}$ with each $x_i \geq 0$, and

$$\|\mathbf{x}\| \leq \frac{1}{\sqrt{n}},$$

by the Cauchy-Schwarz inequality we have

$$x_1 + \cdots + x_n \leq \sqrt{n}\sqrt{x_1^2 + \cdots + x_n^2} \leq 1.$$

Now we define $t_i = x_i/(x_1 + \cdots + x_n)$, and find that

$$t_i \geq 0, t_1 + \cdots + t_n = 1.$$

Further, $\mathbf{x} = (x_1 + \cdots + x_n)(t_1, \cdots, t_n)$, so for such $\mathbf{x}$

$$f(\mathbf{x}) \leq [1 - (x_1 + \cdots + x_n)]f(\mathbf{0}) + (x_1 + \cdots + x_n)f(t_1, \cdots, t_n).$$

This implies that

$$f(\mathbf{x}) - f(\mathbf{0}) \leq (M - f(\mathbf{0}))(x_1 + \cdots + x_n) \to 0 \text{ as } \mathbf{x} \to 0.$$

We can certainly cover $\mathbb{S}^{n-1}$ by a finite number of similarly constructed sets, and carry out this argument, which would allow us to prove that

$$\limsup_{\mathbf{x} \to \mathbf{0}} f(\mathbf{x}) \leq f(0).$$

To prove $\liminf_{\mathbf{x} \to \mathbf{0}} f(\mathbf{x}) \geq f(0)$, we use the property of the secant lines as done in the one-dimensional case. For any $\mathbf{x}$ such that $\|\mathbf{x}\| \leq \frac{1}{\sqrt{n}}$. We bound $f(\mathbf{x})$ from below by the secant line through $(-\frac{\mathbf{x}}{\|\mathbf{x}\|\sqrt{n}}, f(-\frac{\mathbf{x}}{\|\mathbf{x}\|\sqrt{n}}), (\mathbf{0}, f(\mathbf{0}))$:

$$f(\mathbf{x}) - f(\mathbf{0}) \geq \frac{f(\mathbf{0}) - f(-\frac{\mathbf{x}}{\|\mathbf{x}\|\sqrt{n}})}{(\sqrt{n})^{-1}}\|\mathbf{x}\|.$$

Since the slope $\frac{f(\mathbf{0})-f(-\frac{\mathbf{x}}{\|\mathbf{x}\|\sqrt{n}})}{(\sqrt{n})^{-1}}$ has a lower bound due to the upper bound of $f(-\frac{\mathbf{x}}{\|\mathbf{x}\|\sqrt{n}})$, this allows us to conclude that $\liminf_{\mathbf{x}\to\mathbf{0}} f(\mathbf{x}) \geq f(0)$.

The secant line property of a convex function implies that, if $[a, a+\epsilon]$ is in the domain of a convex function, then the slope of the secant line $\frac{f(x)-f(a)}{x-a}$ has a limit as $x \to a+$, although this limit could be $-\infty$. If a is an interior point of the domain, then picking some $c < a$ in the domain implies a lower bound of $\frac{f(x)-f(a)}{x-a}$ in terms of $\frac{f(c)-f(a)}{c-a}$ when $x > a$, so in such a case, f has finite **left derivative** $D_-f(a)$ and **right derivative** $D_+f(a)$ at a, and $D_-f(a) \leq D_+f(a)$. Furthermore, for any $k, D_-f(a) \leq k \leq D_+f(a)$,

$$\frac{f(x)-f(a)}{x-a} \geq D_+f(a) \geq k, \text{ for } x > a;$$
$$\frac{f(x)-f(a)}{x-a} \leq D_-f(a) \leq k, \text{ for } x < a.$$

This then implies that

$$f(x) \geq f(a) + k(x-a) \text{ for all } x \text{ in the domain of } f.$$

Since the right hand side, $f(a) + k(x-a)$, represents a straight line, the above inequality shows that *a convex function has a (linear)* ***support function*** *at any interior point of its domain.*

The support function property of a convex function can be used to give a simple proof of Jensen's inequality.

Theorem 6.4 Jensen's inequality.

Suppose that $f : X \mapsto (A, B)$ and $p(x) \geq 0$ is a density function on X, namely, $\int_X p(x)\,dx = 1$. Suppose that $\phi : (A, B) \mapsto \mathbb{R}$ is convex, then

$$\phi\left(\int_X f(x)p(x)\,dx\right) \leq \int_X \phi(f(x))p(x)\,dx.$$

In words, "ϕ evaluated at the average of f is not more than the average of $\phi \circ f$."

Proof. Set $\bar{f} = \int_X f(x)p(x)\,dx$. It is easy to rule out the possibility that $\bar{f} = A$ or B, so we assume that $A < \bar{f} < B$. Using the support property of ϕ at $\bar{f}$, there exists some k such that

$$\phi(y) \geq \phi(\bar{f}) + k(y - \bar{f}) \text{ for all } y \in (A, B).$$

Substituting y by $f(x)$, multiplying the above inequality by $p(x)$ and integrating over $x \in X$, we get

$$\int_X \phi(f(x))p(x)\,dx \geq \phi(\bar{f})\int_X p(x)\,dx + k\left(\int_X f(x)p(x)\,dx - \bar{f}\int_X p(x)\,dx\right).$$

The right hand side is simply $\phi(\bar{f})$, which proves the Jensen's inequality. ∎

Commonly used cases of Jensen's inequalities include $\phi(y) = -\ln y$ or $y \ln y$. Proofs for Hölder's and Minkowski's inequalities also use convexity in crucial ways.

Checkpoint 6.5 **Prove that** $\ln\left(\int_X e^{u(x)}p(x)\,dx\right) \geq \int_X u(x)p(x)\,dx$ **for** $p(x) \geq 0, \int_X p(x)\,dx = 1$**.**

6.3 Some properties of convex functions of several variables

When proving the continuity of a convex function of several variables, we already saw the complications for multi-dimensions. We do not intend to do a serious study of properties of convex functions of several variables, but only want to briefly discuss a few properties related to the notion of *supporting planes* to illustrate how the notion of compactness comes into play.

Definition 6.6

Let K be a convex set, $\mathbf{x}_0$ be a point on the boundary of K. K is said to have a supporting hyperplane at $\mathbf{x}_0$ if there exists a non-zero vector $\mathbf{n}$ such that

$$(\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{n} \geq 0 \text{ for all } \mathbf{x} \in K.$$

Definition 6.7

Let K be a convex set and $f(\mathbf{x})$ be a convex function defined on K. Let $\mathbf{x}_0 \in K$. The graph of f at $(\mathbf{x}_0, f(\mathbf{x}_0))$ is said to have a supporting hyperplane if there exists a non-zero vector $\mathbf{v}$ such that

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \mathbf{v} \cdot (\mathbf{x} - \mathbf{x}_0) \text{ for all } \mathbf{x} \in K.$$

Note that we use the same terminology in these two contexts, but they have a slight distinction, as illustrated by the simple example $f(x) = -\sqrt{x}$ on $[0, 1]$. As a function it does not have a supporting hyperplane (a straight line here) at $x = 0$, but its epigraph has a supporting hyperplane at $(0, 0)$ (a vertical line).

For a convex function of a single variable we gave a proof of the existence of a supporting line at any interior point using the property of secant lines. We can apply this argument along any direction to a convex function of several variables, but it alone would not give us a supporting hyperplane at a point on the graph. The extension to multi-dimensions would necessarily involve some kind of limiting argument and compactness. We will discuss the following theorems.

Theorem 6.8 Existence of a supporting hyperplane of a closed convex set.

Any boundary point of a closed convex set has a supporting hyperplane.

Proof. Let $\mathbf{x}_0 \in K$ be a boundary point of the closed convex set K. Then there exists a sequence $\mathbf{x}_k \notin K, \mathbf{x}_k \to \mathbf{x}_0$. Each $\mathbf{x}_k$ also has a closest point $\mathbf{p}_k \in K$. This is done either by the Bolzano-Weierstrass compactness theorem or the parallelogram law of the Euclidean norm

$$2\|\frac{\mathbf{p} - \mathbf{q}}{2}\|^2 = \|\mathbf{p} - \mathbf{x}_k\|^2 + \|\mathbf{q} - \mathbf{x}_k\|^2 - 2\|\frac{\mathbf{p} + \mathbf{q}}{2} - \mathbf{x}_k\|^2$$

and the completeness of $\mathbb{R}^n$. This law shows that if $\mathbf{q}_l \in K$ is such that

$$\|\mathbf{q}_l - \mathbf{x}_k\| \to \inf\{\|\mathbf{x} - \mathbf{x}_k\| : \mathbf{x} \in K\},$$

then $\mathbf{q}_l$ is a Cauchy sequence, therefore has a limit.

Next we claim that

$$(\mathbf{p}_k - \mathbf{x}_k) \cdot (\mathbf{x} - \mathbf{p}_k) \geq 0 \text{ for all } \mathbf{x} \in K. \tag{6.1}$$

This follows from considering

$$h(t) := (t\mathbf{x} + (1-t)\mathbf{p}_k - \mathbf{x}_k) \cdot (t\mathbf{x} + (1-t)\mathbf{p}_k - \mathbf{x}_k).$$

Note that $h(t) = \|t\mathbf{x} + (1-t)\mathbf{p}_k - \mathbf{x}_k\|^2$, and $t\mathbf{x} + (1-t)\mathbf{p}_k \in K$ for $0 \le t \le 1$, so $h(0) \le h(t)$ for all $0 \le t \le 1$. It follows that

$$h'(0) = 2(\mathbf{p}_k - \mathbf{x}_k) \cdot (\mathbf{x} - \mathbf{p}_k) \ge 0.$$

Define $\mathbf{n}_k = (\mathbf{p}_k - \mathbf{x}_k)/\|\mathbf{p}_k - \mathbf{x}_k\|$. Then $\mathbf{n}_k$ is a sequence of unit vectors, so there exists a subsequence, still denoted by itself, and a limiting unit vector $\mathbf{n}$ such that $\mathbf{n}_k \to \mathbf{n}$. We also know that $\mathbf{p}_k \to \mathbf{x}_0$ as

$$\|\mathbf{p}_k - \mathbf{x}_0\| \le \|\mathbf{p}_k - \mathbf{x}_k\| + \|\mathbf{x}_k - \mathbf{x}_0\| \le 2\|\mathbf{x}_k - \mathbf{x}_0\|.$$

For each fixed $\mathbf{x} \in K$, dividing through both sides of (6.1) by $\|\mathbf{p}_k - \mathbf{x}_k\|$, and passing to the limit, we get

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0) \ge 0,$$

which is the inequality defining a supporting hyperplane.

In summary, the idea is that, in the absence of a direct construction of a supporting plane at the given point, one finds a relatively easy way to construct a supporting plane at a nearby, but unspecified point, and one then takes a limiting process to obtain a supporting plane at the given point. ■

Theorem 6.9 Existence of a supporting hyperplane of the graph of a convex function at an interior point.

Let $f(\mathbf{x})$ be a convex function defined on the convex set K. If $\mathbf{x}_0 \in K$ in an interior point of K, then the graph of f has a supporting hyperplane at $(\mathbf{x}_0, f(\mathbf{x}_0))$.

Proof. The epigraph $G_f = \{(\mathbf{x}, y) : \mathbf{x} \in K, y \ge f(\mathbf{x})\}$ is a convex set. Its closure $\overline{G_f}$ is a closed convex set, and $(\mathbf{x}_0, f(\mathbf{x}_0))$ is on the boundary of $\overline{G_f}$. By the previous theorem, there exists a non-zero vector $\mathbf{n} = (\mathbf{v}, c)$ such that

$$\mathbf{v} \cdot (\mathbf{x} - \mathbf{x}_0) + c(y - f(\mathbf{x}_0)) \ge 0 \text{ for all } (\mathbf{x}, y) \in \overline{G_f}.$$

Since $\mathbf{x}_0 \in K$ in an interior point of K, we claim that $c \ne 0$. For, otherwise, we would have

$$\mathbf{v} \cdot (\mathbf{x} - \mathbf{x}_0) \ge 0 \text{ for all } \mathbf{x} \in K,$$

which would force $\mathbf{v} = \mathbf{0}$.

Next we claim that $c > 0$. This is because $(\mathbf{x}_0, f(\mathbf{x}_0) + t) \in G_f$ for any $t > 0$, and the above inequality then forces $c > 0$. Now it follows that for any $\mathbf{x} \in K$, applying the above inequality for $y = f(\mathbf{x})$ implies that

$$f(\mathbf{x}) \ge f(\mathbf{x}_0) - c^{-1}\mathbf{v} \cdot (\mathbf{x} - \mathbf{x}_0),$$

which demonstrates a supporting hyperplane to the graph of f at $(\mathbf{x}_0, f(\mathbf{x}_0))$. ■

It is possible to prove this theorem directly using the properties of a convex function, along the lines of proof for the one dimensional case. You should try to construct such a proof, at least for the two dimensional case.

Checkpoint 6.10 The tangent plane of a convex function at a differentiable point is a supporting plane to the graph of the function. Further-

more, if the point is in the interior of the domain, then it is the unique supporting plane.

Hint. If f denotes the function, and $Df(\mathbf{x}_0)$ denotes the gradient of f at $\mathbf{x}_0$, it may be geometrically easier to consider

$$g(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}_0) - Df(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0),$$

which is also convex.

We close this subsection by discussing a more subtle application of convex/concave functions in an optimization problem.

Example 6.11 A different approach to Example 1.11 using concavity.

We introduce the new variables $u_i = a_i^2, v_i = b_i^2$, and reformulate the problem in Example 1.11 in terms of u_i, v_i. The quotient there now becomes

$$\frac{\sum_i \sqrt{u_i v_i}}{\sqrt{(\sum_i u_i)(\sum_i u_i)}},$$

and the constraints become

$$a^2 \leq u_i \leq A^2; b^2 \leq v_i \leq A^2.$$

Our argument will be based on the following observation.

1. $\sqrt{uv}$ is a concave function in the first quadrant.
2. For any (u, v) in the rectangle $[a^2, A^2] \times [b^2, B^2]$, there exists unique (p, q) with $p, q \geq 0$, such that

$$(u, v) = p(a^2, B^2) + q(A^2, b^2).$$

3. In the set up above, we have

$$\sqrt{uv} \geq paB + qAb,$$

with equality iff (u, v) equals (a^2, B^2), or (A^2, b^2), equivalently, $(p, q) = (1, 0)$, or $(0, 1)$.

For the second item, note that for any (u, v) in the rectangle $[a^2, A^2] \times [b^2, B^2]$, there exists a unique $s > 0$, such that $s(u, v)$ lies on the diagonal from (a^2, B^2) to (A^2, b^2), which implies the existence of a unique $0 \leq t \leq 1$ such that

$$s(u, v) = (1 - t)(a^2, B^2) + t(A^2, b^2).$$

This then implies our desired relation.

We remark that in proving the last item above, only the (strict) concavity of $\sqrt{uv}$ along the diagonal from (a^2, B^2) to (A^2, b^2) is used. It is this last item that makes it possible to bound $\sum_i \sqrt{u_i v_i}$ from below.

Now for each (u_i, v_i), we find (p_i, q_i) according to the second item above

$$(u_i, v_i) = p_i(a^2, B^2) + q_i(A^2, b^2),$$

then we can bound the quotient as

$$\frac{\sum_i \sqrt{u_i v_i}}{\sqrt{(\sum_i u_i)(\sum_i u_i)}} \geq \frac{\sum_i (p_i aB + q_i Ab)}{\sqrt{[\sum_i (p_i a^2 + q_i A^2)][\sum_i (p_i B^2 + q_i b^2)]}}.$$

Setting $p = \sum_i p_i, q = \sum_i q_i,\ \alpha = A/a, \beta = B/b$, and after dividing the quotient on the right hand above by ab, it becomes

$$\frac{p\beta + q\alpha}{\sqrt{(p + q\alpha^2)(p\beta^2 + q)}},$$

and now the task is to find the infimum of this quotient when $p, q \geq 0$, and identify when equality can occur. This calculus problem was formulated in Checkpoint 1.12, and the answer is $\frac{2\sqrt{\alpha\beta}}{\alpha\beta+1}$.

Checkpoint 6.12 Prove a lower bound of $\frac{\int_{x_1}^{x_2} f(x)g(x)\,dx}{\sqrt{\left(\int_{x_1}^{x_2} f(x)^2\,dx\right)\left(\int_{x_1}^{x_2} g(x)^2\,dx\right)}}$ **when** $f(x), g(x)$ **are subject to positive upper and lower bounds.** This problem is from #93 in Part II, Chapter 2 of Polya and Szegö's classic "Problems and Theorems in Analysis I"

Let a, A, b, B be positive numbers such that $a < A, b < B$. If the two functions $f(x)$ and $g(x)$ are integrable over the interval $[x_1, x_2]$, and $a \leq f(x) \leq A, b \leq g(x) \leq B$ on the interval. Then

$$\frac{\int_{x_1}^{x_2} f(x)g(x)\,dx}{\sqrt{\left(\int_{x_1}^{x_2} f(x)^2\,dx\right)\left(\int_{x_1}^{x_2} g(x)^2\,dx\right)}} \geq \frac{2}{\sqrt{\frac{AB}{ab} + \sqrt{\frac{ab}{AB}}}}.$$

7 Some Comments on Calculus of Several Variables, including Green's Theorem, Divergence Theorem, and Stokes' Theorem

Professor Ocone's notes contain some review on the Jacobian of a differentiable map of several variables, and the inverse/implicit function theorems in that setting. We will only make some comments about integrals involving several variables.

Multiple (Riemann) integrals can be defined in a similar fashion as the Riemann integrals in the one variable case. Instead of defining integrals only on closed intervals in the one variable case, there are often needs to define integrals over more general sets in multi-dimensions. However, complications arise in extending to multi-dimensions the concept of partitions over general sets; there are even bounded open sets U in $\mathbb{R}^n$ for which $\int_U f(x)dx$ is not well defined in this fashion for f that are continuous over the closure $\bar{U}$ of U.

However, it is not too hard to establish the following two facts:

- If $f(x)$ is continuous and has compact support in $\mathbb{R}^n$, then the Riemann integral $\int_{\mathbb{R}^n} f(x)dx$ is well defined.
- Suppose U is a bounded open set in $\mathbb{R}^n$ such that its boundary ∂U has a finite cover $\cup V_j$, and each $V_j \cap \partial U$ is modeled as the differentiable image of an open set in $\mathbb{R}^k$ for some $k \leq n-1$, namely, $V_j \cap \partial U = \phi_j(W_j)$, where W_j is a bounded open set in $\mathbb{R}^k$, and ϕ_j is Lipschitz over the closure $\bar{W}_j$ of W_j. Then for any $f(x)$ which is continuous over the closure $\bar{U}$ of U, the Riemann integral $\int_U f(x)dx$ is well defined.

Although the cases covered above contain most cases that we normally encounter, Riemann integral has the obvious defect that it is not well defined for a wide enough class of functions and sets; in particular, reasonable limits of integrable functions may not be integrable. This lack of completeness of integrable functions is a major drawback of Riemann integrals: we would rather keep the completeness, but accept that integrals do not have to be defined in the Riemann fashion. Lebesgue's integration theory was developed to overcome this difficulty, and will be the focus of the first semester's graduate analysis course.

However, there are several aspects of the integration theory in the Euclidean space and its submanifolds which can not be swept under the general framework of Lebesgue's integration theory. These include

- Change of variables formula for integrals in several variables.
- Integrals defined on submanifolds in the Euclidean space (such as hyper surfaces).
- Integration of vector fields and more generally differential forms on submanifolds in the Euclidean space (or more generally on abstract manifolds), and relations between such integrals as given by Green's theorem, Divergence theorem, and Stokes' theorem.

I have two separate documents discussing aspects of the above topics, one is named CurDivergence.pdf[1], another named ChangeofVar.pdf[2]. You may consult these documents if you need a review on certain topics. Here we will discuss some examples involving these aspects.

Example 7.1 (January 2016 WQ).

Consider the vector field $\mathbf{F}(x,y,z) = (-4xz^3, 0, z^4)$ in $\mathbb{R}^3$ and let S be the (compact) portion of the paraboloid $z = x^2 + y^2$ having $z \leq 9$. Use Stokes' theorem to evaluate

$$\int_S \mathbf{F}(x,y,z) \cdot d\mathbf{S},$$

where $d\mathbf{S}$ is the vector surface element corresponding to the upward pointing normal vector.

Solution. This integral could have been evaluated directly using the definition. One first recognizes that S can be parametrized as a graph over $\{(x,y) : x^2 + y^2 \leq 9\}$, so

$$\int_S \mathbf{F}(x,y,z) \cdot d\mathbf{S} = \int_{x^2+y^2\leq 9} \mathbf{F}(x,y,z)\text{ů}(-z_x, -z_y, 1)\, dx\, dy,$$

where we have used

$$\frac{(-z_x, -z_y, 1)}{\sqrt{z_x^2 + z_y^2 + 1}} \text{ as the unit normal of } S,$$

and $\sqrt{z_x^2 + z_y^2 + 1}\, dx\, dy$ as the area element of S.

However, the instruction asks for using Stokes' theorem, which suggests that we need to recognize $\mathbf{F}(x,y,z)$ to be equal to or related to the curl of some vector field. Using the differential form formulation for such integrals, we associate $\mathbf{F}(x,y,z)$ with the differential form

$$\mathbf{F}_1(x,y,z)\, dy \wedge dz + \mathbf{F}_2(x,y,z)\, dz \wedge dx + \mathbf{F}_3(x,y,z)\, dx \wedge dy$$

[1] https://rutgers.box.com/s/qhg3yu17o4tka6xikxwo1xshp0nm5b2r
[2] https://rutgers.box.com/s/ur2mhmgqq7f61jll1qw6pz3z72g7b9ub

$$
\begin{aligned}
&= -4xz^3\,dy \wedge dz + z^4\,dx \wedge dy \\
&= d(xz^4) \wedge dy = d(xz^4\,dy),
\end{aligned}
$$

so

$$\int_S \mathbf{F}(x,y,z) \cdot d\mathbf{S} = \int_S d(xz^4\,dy) = \int_{\partial S} xz^4\,dy,$$

by the Stokes' theorem. On ∂S, $z = 9$, so

$$\int_{\partial S} xz^4\,dy = 9^4 \int_{\partial S} xdy = 9^4 3^2 \pi,$$

where we have used $\int_{\partial S} xdy$ to be the area enclosed by the circle $x^2 + y^2 = 9$ (This can be seen by Stokes' theorem again: $\int_{\partial S} x\,dy = \int_{x^2+y^2 \le 9} dx \wedge dy = 3^2\pi$).

One could also apply the Divergence Theorem to evaluate this integral by recognizing the S, together with the top $T := \{(x, y, 9) : x^2 + y^2 \le 9$ encloses a solid region V, and choosing $(0, 0, 1)$ as the unit normal to T exterior to V, we see that

$$\int_T \mathbf{F} \cdot (0,0,1)\,dx\,dy - \int_S \mathbf{F}(x,y,z) \cdot d\mathbf{S} = \int_V \operatorname{div} \mathbf{F}\,dx\,dy\,dz.$$

But div $\mathbf{F} = (-4xz^3)_x + 0 + (z^4)_z = 0$, so

$$\int_S \mathbf{F}(x,y,z) \cdot d\mathbf{S} = \int_T \mathbf{F} \cdot (0,0,1)\,dx\,dy = \int_T z^4\,dx\,dy = 9^4 3^2 \pi.$$

Checkpoint 7.2 Let $\vec{F}$ denote the vector field $\vec{F}(\vec{x}) = (x, y, -2z)$ in $\mathbb{R}^3$, and S_R denote the upper hemisphere $\left\{(x,y,z) : x^2 + y^2 + z^2 = R^2, z \ge 0\right\}$. Evaluate the integral

$$\int_{S_R} \vec{F}(\vec{x}) \cdot \vec{n}(\vec{x})\,d\sigma(\vec{x}),$$

where $\vec{n}(\vec{x})$ denotes the unit normal vector to S_R at $\vec{x} \in S_R$, pointing upward, and $d\sigma(\vec{x})$ denotes the area element of S_R.

Example 7.3 (January 2007 WQ).

Suppose that Ω is a bounded domain in $\mathbb{R}^3$ whose boundary, $\partial\Omega$, is a C^1 hypersurface. Let $\nu(\mathbf{y}) = (\nu_1(\mathbf{y}), \nu_2(\mathbf{y}), \nu_3(\mathbf{y}))$ denote the unit exterior normal vector to $\partial\Omega$ at $\mathbf{y} \in \partial\Omega$, and $d\sigma(\mathbf{y})$ denote the area form for $\partial\Omega$.

(a). Prove that

$$\int_{\partial\Omega} \frac{\mathbf{y} \cdot \nu(\mathbf{y})}{|\mathbf{y}|^3}\,d\sigma(\mathbf{y}) = \begin{cases} 0, & \text{if } 0 \in \mathbb{R}^3 \setminus \overline{\Omega}; \\ 4\pi, & \text{if } 0 \in \Omega. \end{cases}$$

(b). Fix a domain Ω satisfying the assumptions above, define for $\mathbf{x} \in \mathbb{R}^3$

$$V_i(\mathbf{x}) = \int_{\partial\Omega} |\mathbf{x} - \mathbf{y}|^{-1} \nu_i(\mathbf{y})\,d\sigma(\mathbf{y}).$$

Prove that

$$\sum_{i=1}^{3} \frac{\partial V_i(\mathbf{x})}{\partial \mathbf{x}^i} = \begin{cases} 0, & \text{if } \mathbf{x} \in \mathbb{R}^3 \setminus \overline{\Omega}; \\ 4\pi, & \text{if } \mathbf{x} \in \Omega. \end{cases}$$

Solution. The domain Ω is not given explicitly, so we should not expect to evaluate the integral directly. Since $\partial\Omega$ is a closed surface, the integral formulated is in a form that the Divergence Theorem may be applicable. The vector field on $\partial\Omega$, $\frac{\mathbf{y}}{|\mathbf{y}|^3}$ has a smooth extension inside of Ω when $0 \in \mathbb{R}^3 \setminus \overline{\Omega}$, in which case a direct application of the Divergence Theorem gives

$$\int_{\partial\Omega} \frac{\mathbf{y}\cdot\nu(\mathbf{y})}{|\mathbf{y}|^3}\, d\sigma(\mathbf{y}) = \int_{\Omega} \operatorname{div}\left(\frac{\mathbf{y}}{|\mathbf{y}|^3}\right) d\mathbf{y} = 0,$$

as $\operatorname{div}\left(\frac{\mathbf{y}}{|\mathbf{y}|^3}\right) d\mathbf{y} = 0$.

When $0 \in \Omega$, the extension inside $\frac{\mathbf{y}}{|\mathbf{y}|^3}$ becomes singular at $\mathbf{y} = \mathbf{0}$, so we can't apply the Divergence Theorem directly inside of Ω. We need to take a small $\epsilon > 0$ so that the ball $\overline{B_\epsilon(\mathbf{0})} \subset \Omega$, and we can apply the Divergence Theorem on $\Omega \setminus \overline{B_\epsilon(\mathbf{0})}$ to get

$$\begin{aligned}&\left(\int_{\partial\Omega} - \int_{\partial B_\epsilon(\mathbf{0})}\right) \frac{\mathbf{y}\cdot\nu(\mathbf{y})}{|\mathbf{y}|^3}\, d\sigma(\mathbf{y}) \\ &= \int_{\Omega\setminus\overline{B_\epsilon(\mathbf{0})}} \operatorname{div}\left(\frac{\mathbf{y}}{|\mathbf{y}|^3}\right) dy = 0.\end{aligned}$$

Using $|\mathbf{y}| = \epsilon$ on $\partial B_\epsilon(\mathbf{0})$, we get

$$\int_{\partial B_\epsilon(\mathbf{0})} \frac{\mathbf{y}\cdot\nu(\mathbf{y})}{|\mathbf{y}|^3} d\sigma(\mathbf{y}) = \int_{\partial B_\epsilon(\mathbf{0})} \epsilon^{-2}\, d\sigma(\mathbf{y}) = 4\pi,$$

and conclude that

$$\int_{\partial\Omega} \frac{\mathbf{y}\cdot\nu(\mathbf{y})}{|\mathbf{y}|^3}\, d\sigma(\mathbf{y}) = 4\pi,$$

in this case.

Example 7.4 (January 2019 WQ).

Let $(P(x,y), Q(x,y))$ be a C^1 vectorfield in $\mathbb{R}^2 \setminus \{(0,0)\}$. It is said to be curl free in $\mathbb{R}^2 \setminus \{(0,0)\}$ if $\partial_x Q(x,y) - \partial_y P(x,y) = 0$ there; it is said to have a potential function in $\mathbb{R}^2 \setminus \{(0,0)\}$ if there exists a C^2 function $\phi(x,y)$ in that region such that $(P(x,y), Q(x,y)) = (\partial_x\phi(x,y), \partial_y\phi(x,y))$ there.

(a). Prove that $(P(x,y), Q(x,y))$ is curl free in $\mathbb{R}^2 \setminus \{(0,0)\}$ iff for any $(x_0,y_0) \in \mathbb{R}^2 \setminus \{(0,0)\}$, $\int_{\partial D_r(x_0,y_0)} P(x,y)\, dx + Q(x,y)\, dy = 0$ for all$0 < r < \sqrt{x_0^2 + y_0^2}$, where $D_r(x_0,y_0)$ is the disc of radius r centered at (x_0,y_0).

(b). Prove that for any curl free vectorfield $(P(x,y), Q(x,y))$ in $\mathbb{R}^2 \setminus \{(0,0)\}$, there exists a unique $c \in \mathbb{R}$ such that

$$(P(x,y), Q(x,y)) - c(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2})$$

has a potential function in $\mathbb{R}^2 \setminus \{(0,0)\}$.

Solution. (a). Applying Green's Theorem, we have

$$\int_{\partial D_r(x_0,y_0)} P(x,y)\, dx + Q(x,y)\, dy = \int\int_{D_r(x_0,y_0)} (\partial_x Q(x,y) - \partial_y P(x,y))\, dx\, dy.$$

The conclusion follows from this.

(b). First, note that, for any $0 < r_1 < r_2$,

$$\int_{\partial D_{r_2}(0,0)} P(x,y)\,dx + Q(x,y)\,dy - \int_{\partial D_{r_1}(0,0)} P(x,y)\,dx + Q(x,y)\,dy$$
$$= \int\int_{r_1^2 \le x^2+y^2 \le r_2^2} \left(\partial_x Q(x,y) - \partial_y P(x,y)\right)\,dx\,dy = 0,$$

so $\int_{\partial D_r(0,0)} P(x,y)\,dx + Q(x,y)\,dy = 2\pi c$ is independent of $0 < r$ for some constant $c \in \mathbb{R}$. Note also that $(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2})$ is also curl free in $\mathbb{R}^2 \setminus \{(0,0)\}$, and

$$\int_{\partial D_r(0,0)} \left(-\frac{y dx}{x^2+y^2} + \frac{x dy}{x^2+y^2}\right) = 2\pi$$

is also independent of $0 < r$. Setting

$$(\widehat{P}(x,y), \widehat{Q}(x,y)) = (P(x,y), Q(x,y)) - c(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}),$$

then $\int_{\partial D_r(0,0)} \widehat{P}(x,y)\,dx + \widehat{Q}(x,y)\,dy = 0$ for all $0 < r$. In the simply connected region $U \subset \mathbb{R}^2 \setminus \{(0,0)\}$ by deleting the negative x-axis, by Poincarè Lemma, $(\widehat{P}(x,y), \widehat{Q}(x,y))$ has a well defined potential function $\phi(x,y)$, and

$$\phi(x,y) = \int_{C_{(x,y)}} \widehat{P}(x,y)\,dx + \widehat{Q}(x,y)\,dy,$$

where $C_{(x,y)}$ is the arc of circle from $(\sqrt{x^2+y^2}, 0)$ to (x,y). Due to the property that $\int_{\partial D_r(0,0)} \widehat{P}(x,y)\,dx + \widehat{Q}(x,y)\,dy = 0$ for all $0 < r$, $\phi(x,y)$ extends to be a continuous function on $\mathbb{R}^2 \setminus \{(0,0)\}$. To see that the extended ϕ is C^2 in a neighborhood of any point $(x,0)$ along the negative x-axis, note that each such point has a small disc $D \subset \mathbb{R}^2 \setminus \{(0,0)\}$ on which $(\widehat{P}(x,y), \widehat{Q}(x,y))$ has a well defined C^2 potential function ψ. Since ψ and ϕ differ by a constant in this disc, it follows that ϕ is also C^2 in this disc.