分析導論

An Introduction to Mathematical Analysis

Ching-hsiao Arthur Cheng 鄭經斅

Copyright

Contents

第一版序

這份高等微積分（或說基礎分析）教材能完成，必須感謝國立中央大學（以下簡稱中大）數學系讓我連續三年教高等微積分（以及國立清華大學數學系的壓力？）。

長久以來，高等微積分（以下簡稱高微）該教什麼內容有著一些爭論，有些人認為既然是高微，就是把微積分所有的東西嚴謹化，淺嚐到此即可，但是也有一派人－包含我自己在內，認為高微就是分析導論，認為除了微積分的嚴謹化之外，還需要加入一些分析的基礎訓練，例如抽象賦距空間（metric space）的點拓（point-set topology）理論，以及像是均勻收斂（uniform convergence）相關的極限交換理論、Arzelà-Ascoli 定理、Stone-Weierstrass 定理等等，都該是此課程必備的內容。近年來也陸續有學校將高微這門課正名為分析導論，而我也認為中大數學的學生程度還能接受較抽象並嚴格的數學基礎訓練，所以經過思考後，還是決定以「分析導論」為教授此門課的主要內容。

決定好要教「高微」還是「分析導論」後，就必須掙扎要選擇什麼樣的教材了。先看一些較新（意思是還沒絕版、買得到）的書吧！近年來 Wade 的 An Introduction to Analysis 也變成高微教科書的常客了。這本書的優點是把微積分的理論交代得非常完整且清楚，並交代了在實數空間的點拓這個框架下面均勻收斂的部份相關理論。唯二的缺憾是抽象的部份只有在書的中段加入的一般賦距空間中的點拓，以及並沒有講到 Arzelà-Ascoli 這個在進階分析學習中常常被假設已知的重要定理，對「分析導論」來說我認為並不是那麼足夠。而 Folland 的 Advanced Calculus，顧名思義就是高微了，所以很快就被淘汰出選擇的名單了。

接下來談談 Rudin 與 Apostol 所著的兩本曾是我當學生那個年代常用的高微教科書。Rudin 的 Principles of Mathematical Analysis 這本書我個人覺得一來過於精簡不適合初學者，二來在積分理論的鋪陳上直接省略了 Riemann 積分裡面多重積分的部份（單變數的部份由 Riemann-Stieltjes 積分取代），導致重要的一些定理，如 Fubini 定理和積分變數變換的公式整個被省略（雖然有用微分形式的積分講了向量微積分裡面的重要定理），

因此雖說大家都說 Rudin 的書是高微的聖經，我個人還是傾向不使用 Rudin 的書當高微教科書；而 Apostol 的 Mathematical Analysis，在單變數微積分的理論為止我覺得寫得很不錯，但在進入多變數之後整本書的內容開始變得比較雜亂，沒有明確的目標與重點，加上內容多到我個人不覺得可以在一學年內教完，所以從給學生一本可讀或老師可用的教科書這個角度來說，我捨棄了 Apostol 的書當教材。此外，值得一提的事是，我翻遍了這兩本書，發現並沒有交代 Compactness、Sequential Compactness 以及 Completeness + Total Boundedness 這三者在賦距空間裡面的等價性，這個在我現在看來形同「點拓基本定理」的東西，如果教材裡完全沒提及，著實不容易讓人接受。

另外有一本 Marsden & Hoffman 的 Elementary Classical Analysis，在某段時間裡也算國內外常用的高微／分析教科書。這本書選材算難易適中，編排順序上我覺得很不錯，但是證明在進入多變數章節之後嚴謹性不足，這對要學習嚴謹證明的數學系學生來說，拿來當教材又是一個災難（但對於只想知道定理的學生而言說也算是足夠了）。

最後講講我學生時代修高微時的“官定”教科書：Protter & Morrey 的 A First Course in Real Analysis 這本書。將“官定”加了引號，是因為當年事實上教我們高微的「大雄」老師並沒有真的照這本書教，例如當時一開始也是像 Rudin 或 Apostol 的書所排的順序一樣先教了賦距空間中的點拓（而 P&M 這本書是在第六章才講點拓），也因此我的點拓實際上是從 Apostol 的書學來的。然而，如果這本書順序編排得不好的話為什麼要選它當教科書呢？我個人的觀察是，這本書的亮點是在講如何對瑕積分定義的函數做微分（在 P&M 這本書的第 11 章 Functions Defined by Integrals; Improper Integrals）！由於沒有實分析（Real Analysis）中幾個重要的收斂定理的輔助，這本書盡了最大的可能介紹在什麼條件之下怎麼對瑕積分所定義出來的函數微分。上述的幾本書中只有 Marsden & Hoffman 的書有提到這方面的理論，但也只是很輕描淡寫地帶過而已，所以 P&M 的書是上面提到的幾本書裡，在這部份講得最詳細的一本。這部份是否需要列入基礎分析見仁見智，不過我當年的確在這裡搞得焦頭爛額。附帶一提，我們當年使用 P&M 這本書也不是整本都上，有時候還是會跳到別的高微書去上些別的內容，例如 Morse Index。

講了這麼多高微教材，那在我心中一本好的高微（或基礎分析）教科書，應該具備有哪些要素呢？我認為該有的要素如下：

1. 相對完整的點集拓樸理論（亦即包含了「點拓基本定理」)；
2. 相對完整的定義在一般賦距空間上的連續函數的相關理論；
3. 相對完整的單變數與多變數微分與 Riemann 積分理論（但不觸及 Lebesgue 積分）；

4. 容易理解的函數列收斂相關理論；
5. 可供練習而且能由淺入深的習題；
6. 可行的一學年可教授完畢的教材。

上面這幾點堅持，加上近年來平板電腦的普及，使得我也希望能提供一份直接能在電腦上方便學生閱讀與操作的教材，這些理由都催生了這份教材。

另外由於教學的需要，在寫下這份教材的同時，我也思考著要怎麼讓學生更容易接受一些抽象的定義和抽象的定理，因此在這份教材上可能偶爾會有（後來被學生形容為好具體的）驚人之作。例如在講數列極限的定義時，雖然定義還是以 ε-N 的方式呈現，但是也加註了一個等價的定義：一個數列 $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ 被說是收斂到 x 如果

$$\forall \varepsilon > 0, \#\Big\{n \in \mathbb{N} \,\Big|\, x_n \notin (x-\varepsilon, x+\varepsilon)\Big\} < \infty\,.$$

教學實驗後的結果顯示，用數集合裡面點的個數這個方式來講極限，會比抽象的 ε-N 來得容易讓學生接受，而且（不收斂的）反敘述也好寫許多。所以怎麼用數集合裡面點的個數這個概念來講 $\limsup$ 和 $\liminf$ 這兩個主題，也出現在這份教材裡。當然，ε-N 這個抽象的概念最後還是得被引進（為了之後講函數數列的收斂），但是在讓學生容易接受的前題之下，我認為這是一個很好的嘗試。

接著，在講到 Arzelà-Ascoli 定理這個部份時，絕大多數提及此定理的書似乎都只注重在直接證明這個結果，沒交代太多來龍去脈。我在教到這邊時，也思考著怎麼讓學生更容易接受為什麼會有或是會需要這個定理。因此我從研究均勻收斂和逐點收斂（pointwise convergence）的差別出發，先證明了定義在緊緻集（compact set）上逐點收斂的連續函數列，

均勻收斂若且唯若等度連續（equi-continuous）

這個性質（Corollary B.7 和 Lemma B.9），再從這個性質出發去探討 diagonal process 與逐點收斂間的關聯性（Lemma B.11），最後再敘述完整的 Arzelà-Ascoli 定理。我發現這樣的 approach 會讓整個 Arzelà-Ascoli 定理來得更容易 accessible。

接著是多變數函數微分的部份。目前我看到的書，並沒有特別講從一個賦範向量空間（normed vector space）映到另一個賦範向量空間的 map 要怎麼微分，都是講從 $\mathbb{R}^n$ 映到 $\mathbb{R}^m$ 的函數的微分，然後講到高次微分時就含糊其詞，有些書是只講高次偏微分帶過，有些書就直接省略不碰觸這個議題。由於在賦範向量空間之間的 map，其微分的概念可

以用來講 Fréchet 微分，因此我特別在講多變數函數的微分這部份時一開始就採取相對較抽象的作法，好處是這在講高次微分時還是可以講得通（因為微分之後所形成的 map 是落在一個賦範空間裡），但是壞處是為了充分了解這個概念，還是得要講一下 bounded linear map 的 operator norm 以及相關的理論。但是總體言之，我認為這種講法瑕不掩瑜，因為在觀念的鋪陳上更為完整，而且能為將來學 Fréchet 微分鋪路，算是一舉數得。

然後，多數教材裡的反函數定理及隱函數定理，都著重在證明局部的（local）反函數定理，並用反函數定理證明隱函數定理，對於全域的（global）反函數定理（我暱稱其為反函數「大」定理）並沒有任何著墨，而全域的反函數定理在判斷一個 map 是不是 diffeomorphism 算是個重要的方法，所以這份教材也敘述並證明了全域的反函數定理。另外，隱函數定理獨立於反函數定理的證明，對於如何解 under-determined 非線性系統會起到提示的作用，因此我也特別花了一些篇幅直接證明了隱函數定理。這些努力除了想讓這份教材變得更與眾不同外，也希望提供更多更實際有用的定理與證明過程。

另外，在基礎分析的學習中，關於積分理論的鋪陳上 Lebesgue 定理（Theorem 6.32）也常被忽略，但從積分理論的「完整性」來說，我覺得 Lebesgue 定理是有非常畫龍點睛的效果的。例如學過實分析（Real Analysis）的人都會知道 Lebesgue 積分是定義在 Lebesgue 可測函數上面，然後積分區城一般來說也都是 Lebesgue 可測，而在學習 Riemann 或是瑕積分時，我們事實上也可以有類似的看法：若定義

1. “Riemann 可測” 集（不一定要有界）是邊界測度為零的集合；

2. “Riemann 可測” 函數是其不連續點形成的集合測度為零的函數；

那麼所謂的 Riemann 積分與瑕積分，就能在這樣的框架設定之下去進行。不過由於並沒有真正的 “Riemann 可測” 這個概念（對有界集合來說上述的 “Riemann 可測” 等價於 Jordan 可測這個已有的概念，但是對 unbounded set 這個瑕積分會關注的情況來說都沒有人做特別的著墨），所以我們在講到瑕積分時定理的敘述上常常會顯得比較囉嗦，例如在設定上一定得要假設集合 A 的邊界測度為零、函數 f 的不連續點所形成的集合測度為零等等，但是讀者對上述的 “Riemann 可測” 的看法了然於心的話，那在看本教材瑕積分相關定理時只要把那些設定的敘述都轉換成 “Riemann 可測”，馬上就能切入實際的條件和結果。這份教材在講到瑕積分時，就是從 “Riemann 可測” 這個觀點出發進一步做理論鋪陳，相信會讓人耳目一新。

本教材的另一個亮點則是盡可能地在 Riemann 積分及瑕積分的框架下，講了 Monotone Convergence Theorem 以及 Bounded Convergence Theorem（或其瑕積分的對應

版 Dominated Convergence Theorem)，這可以讓並沒打算學實分析的讀者，讀完這份教材後便有這幾個重要的函數列積分收斂定理可以使用。這幾個定理分布在 §6.5(Corollary 6.67 和 Corollary 6.68)、§6.4(Theorem 6.69 和 Theorem 6.70)還有 §6.6(Theorem 6.102 和 Theorem 6.103)，其中 §6.5 只講這些定理的單變數 Riemann 積分版、§6.4 講的是這些定理的 Riemann 多重積分版，而 §6.6 則是將這些大定理在不討論 Lebesgue 積分下的最佳化版本(亦即瑕積分版)也給了證明。一旦有了這些定理，學習由(瑕)積分所定義的函數如何做微分就相對上容易地多，不需要再因為缺乏工具而需要像 P&M 的書裡必須對「均勻收斂」(另一個均勻收斂的概念，所以加了引號)做文章。

由於調和分析(Harmonic Analysis)的基礎是 Fourier series 與 Fourier transform，所以不能免俗地我還是在本教材的最後一章交代了 Fourier 的理論。在中大這幾年中曾經有位外系的學生來問我 Fourier series 和 Fourier transform 之間的關係，他說很多老師說明兩者的關係時，就直接解釋「在有限空間是以級數的方式呈現，在無窮的空間就把級數改成積分」，但這樣的解釋無法說服他，所以特地來數學系尋求一個說法。被外系的學生問這個問題，實在不能搬出一大套數學理論說服他就是如此這般所以 Fourier transform 是對的，而是要奠基於學生對 Fourier series 的信任去講為什麼 Fourier transform 成立。為了能好好將這兩件事連結起來，最後我找到一個尚可接受的說法，我相信這個說法對外系的學生來說，應該是可以接受的，這個說法也一併呈現在 372 頁一開始的部份。

介紹完了這份教材的特色之後，我想小談一下為什麼我並不想在這份教材中加入 Lebesgue 積分的理論。事實上，由於 Lebesgue 積分在理論的鋪陳上幾乎可以取代 Riemann 積分，很多老師對於要不要教 Riemann 積分的理論會有所遲疑，一個比較“實際”的論點是『學了實分析之後 Riemann 積分的理論就過時了，那為何還要花時間教 Riemann 積分？』，我個人並不是很同意這樣的觀點，原因有以下幾點：

1. 最根本的原因是多數的 Lebesgue 積分的計算都是換成 Riemann 積分利用變數變換公式、Fubini 定理再加上微積分基本定理去算出來的。Lebesgue 積分在抽象理論發展上也許較 powerful，但是在實際計算上還是得回到這些用 Riemann 積分發展出來的工具去做計算(我稍微想了一下，變數變換公式要用 Lebesgue 積分的理論直接證明出來似乎有點天方夜譚)。

2. 因為 Lebesgue 積分的理論奠基於測度論之上，當測度論沒好好講清楚之前，Lebesgue 積分就無從談起，在積分理論的教學方面並不是那麼友善。此外，Riemann 積分和瑕積分相對上較容易上手也更直觀，教學時會容易得多，實際應用時也是以 Riemann 積分和瑕積分為主。

3. 只要有了 Riemann 積分版的 Monotone Convergence Theorem 和 Dominated Convergence Theorem ，在實分析裡 Lebesgue 積分的優勢只剩下建構 L^p 這個完備空間。但是在很多應用的場合中，L^p 空間的完備性較少被使用，而 L^p-norm 所滿足的 Hölder 不等式及 Minkowski 不等式都能在 Riemann 積分的框架下被證明出。

基於以上的幾點原因，我相信這份在積分理論上有著 Fubini Theorem、Tonelli Theorem、Change of Variable Formula、Monotone Convergence Theorem 和 Dominated Convergence Theorem 的教材，會滿足絕大多數人的需求。

最後要說說本教材不足之處。由於考量到一學年的教學需要，我刻意地捨棄了向量微積分（其中包含 Green、Stokes 和 divergence 三個與向量場相關的微積分定理）這個單元。不是這個部份不重要，而是從訓練基礎分析的觀點，向量微積分的內容反而沒有學習的急迫性，加上為了講清楚向量分析，一些曲線、曲面的幾何相關理論是避不開的，所以取捨後還是決定不把這個部份加入本教材。不過我因為本身的研究，在向量微積分方面有著與一般高微書的內容不同的見解，所以我將會在開設向量分析課時，將我的見解弄成另一份專屬於這門課的教材，屆時有興趣的讀者可以再參閱我所寫的向量分析。

鄭經斅 2015 July 19 於中大

第二版序

這份分析導論的教材，是基於之前在國立中央大學教了數年高等微積分後所編之基礎分析教材（俗稱高微共筆，以下以高微共筆稱呼之前的基礎分析教材以資區別），依照特別需要所重新選材編輯而成。

國立中央大學於一零八學年度起分設「計算與資料科學組」（下稱計資組）。計資組的課程經過大幅度地更動，設計的課程皆為了讓學生於大三時能開始進入科學計算、數學影像處理、金融數學以及資料科學這四個領域所準備，其中分析導論課為所有大二必修課如機率導論、統計學、數學建模、數值線性代數、數值微分方程、最佳化方法與應用等課程的分析基礎。

然而教材的編排並不容易，主要都是與其它課程搭配的問題，分成以下兩個不同的方面說明。第一個方面是進度上與其它課程的搭配。例如數值線性代數的課程被安排在大二上學期與分析導論課同時進行，那麼很多數值線性代數中用來求解 $Ax = b$ 的疊代法證明所需要的概念（如賦範空間、矩陣的範數以及判斷收斂性的一般法則），必須在數值線代課教到疊代法之前就先在分析導論課中鋪陳到，因此在介紹實數完備性時，建構實數系這個浩大的工程就不適合出現在課程內容中，以避免拖慢進入賦範空間相關內容的時程；而在講到賦範空間時也必須同時把矩陣範數的性質做比較前期的梳理，讓學生在不久後的數值線代課中不至於因為這些相對較抽象的概念影響了疊代法求解的演算法之學習。再例如同樣被安排在上學期教授的數學建模課中最後的一部份是與變分學相關的內容，而變分的原理主要就是基於 Fréchet 微分的概念，因此若在分析課在上學期就能將定義在賦範空間上的函數之可微性做個梳理，那學生學到變分概念時就更容易接受這些知識。而在上學期將與可微相關的內容結束掉也有很多好處：一來可以在下學期專注在發展 Fourier 分析理論，二來對於安排在下學期的最佳化方法與應用課所需的理論基礎做了最基本的鋪墊。

第二個方面是內容上與其它課程的搭配問題，例如反函數定理與隱函數定理，它們

雖然在分析中是函數可微性非常重要的環節，但在（數學系大學部課程的）應用上通常只會出現在微分幾何，那麼為了在有限的課程中盡量涵蓋其它課程所需要的基本分析知識，反函數、隱函數定理就得忍痛割愛，而高等微積分中原本會講授的許多抽象內容，同樣也必須被修改成較為具體的實用版本或捨棄（可放入習題或是附錄中）。綜合以上兩方面的考量，計資組課程的分析導論最終只能選擇（對之前的高微共筆進行魔改以重新）自編教材。

在盡量維持先前的高微共筆精神的前提下，新版的分析導論主要有以下幾個特點：

1. 針對應用的需求重新編排；
2. 更為精簡的點集拓樸內容；
3. 更為完整的 Fourier 理論；
4. 關於 Fourier 理論的簡單應用。

在此也特別針對第二點「更為精簡的點集拓樸內容」做個簡單的說明。由於在應用上常常需要考量的都是（由演算法生成的）數列之收斂性，因此在講述點集拓樸時，本版分析導論主要想聚焦在「如何判斷一個數列是否收斂或是有無收斂子數列」這件事上，因而捨棄了之前（一般教科書也會採取的）先定義開集合的方式，改採先定義閉集合（是包含所有的 limit points 的集合）然後再展開其它的內容。同樣的，sequential compactness 這件事，也是由 Bolzano-Weierstrass 定理的觀點出發，去問一般的 metric space 下面是否有什麼「性質」，可以保證「滿足該性質的數列」或是「滿足該性質的集合裡的數列」都有收斂子數列（這件事最後在 Problem 3.35 做了完結），這些安排，都是為了讓學生能更熟練與數列收斂相關的性質。而最後在講到 compactness 時，高微共筆中有許多例如以 open cover 定義 compactness 後會先鋪陳一些諸如 compact set 必 closed 等的基本性質，之前關於這方面的鋪陳主要是給學生從不同的角度去理解在 open cover 的 compactness 定義下能走多遠，比較偏向累積抽象定義的熟悉度與持續純數理邏輯的訓練，在本版分析導論中則是在以 open cover 定義 compactness 後直接證明 sequential compactness 與 compactness 的等價性（也因此那些性質就不證自明），並把一些諸如 finite intersection property 與 nested set property 放到習題去來節省授課時間。

最後聊一下 Fourier 理論與應用這部份。理論的部份主要是把 tempered distribution 的 Fourier transform 做了更完整的鋪陳，但是從實際教學的經驗來看，這個部份學生的接受度非常低，可能因為成熟度還不足以應付廣義函數這個想法，但是從為了講 Dirac delta 函數的嚴謹度來說又不得不講，事實上頗令人為難。在應用方面選擇了一般數學系的老師比較少接觸的訊號處理，以及較多老師會接觸的熱方程求解。熱方程求解的部份就不再贅述，而訊號處理的部份主要介紹了「嚴謹版」的 Shannon 取樣定理 (Theorem 10.3)，這個嚴謹版的定理必須要求訊號本身是平方可積的且訊號之 Fourier transform 有 compact support，但是在現實中我們也期待取樣定理對於例如單頻的正弦函數能夠成立。過去幾年我一直設法弱化取樣定理成立的條件，例如訊號是有界連續函數且其 Fourier transform (這是個 tempered distribution) 有 compact support⋯但是一直未能成功，一般文獻也沒有相關定理的證明，因此最後在本教材中呈現的定理還是只有最簡單的取樣定理。這個部份等待有朝一日我能完成這個證明或是等有緣人看到這個需求也願意提供解決之道再進行改寫了。

鄭經斅 2023 Jun. 27 於中大

第三版修訂說明

本版的修訂幅度不大，主要是在原先第二版的內容中新增 Universal Approximation Theorem 與 Kolmogorov–Arnold Representation Theorem 兩節（Section 7.6, 7.7），並配合新增若干習題。加入這兩個主題，一方面是因應近年人工智慧與機器學習的發展，另一方面也是希望從現代的觀點補充本教材原有的函數逼近內容。

函數的表示與逼近原本就是分析學中的基本問題。從 Stone–Weierstrass Theorem、Fourier Analysis，到本教材原先所介紹的相關內容，其核心問題之一都是如何利用具有較簡單結構的函數或函數系統，描述或逼近所研究的函數。近年來神經網路在人工智慧與機器學習中的廣泛應用，使得這些古典問題又在新的脈絡下受到重視；其中，Universal Approximation Theorem 正好提供了一個連結基礎分析與神經網路理論的自然例子。

本版加入 Universal Approximation Theorem，主要是希望說明具有適當結構的神經網路，在什麼意義下可以任意逼近相當廣泛的函數類別；而 Kolmogorov–Arnold Representation Theorem 則從另一個方向說明，多變數連續函數可以利用具有特殊結構的單變數連續函數作精確表示。前者是一個逼近定理，後者則是一個表示定理，兩者的問題意識、數學內容與適用方式並不相同。不過將它們放在一起考察，可以讓讀者從不同角度看見「函數如何被表示」、「函數如何被逼近」以及「模型的結構如何影響其表現能力」這幾個彼此相關的問題。

這次補充的目的，並不是改變本教材原有的課程架構，而是希望讓讀者進一步看見，基礎分析中一些看似古典的問題，如何自然地延伸到當代數學與計算科學所關心的方向。對於有興趣進一步學習神經網路、深度學習或相關數值方法的讀者而言，這兩節也可以作為理解其數學基礎的一個入口。

鄭經斅 2026 Sept. 7 於中大

Chapter 0

Review of Contents from Basic Mathematics

0.1 Sets

Definition 0.1. A ***set*** is a collection of objects called ***elements*** or ***members*** of the set. To denote a set, we make a complete list $\{x_1, x_2, \cdots, x_N\}$ or use the notation

$$\{x : P(x)\} \quad \text{or} \quad \{x \,|\, P(x)\},$$

where the sentence $P(x)$ describes the property that defines the set. A set A is said to be a ***subset*** of S if every member of A is also a member of S. We write $x \in A$ (or A contains x) if x is a member of A, and write $A \subseteq S$ (or S includes A) if A is a subset of S. The empty set, denoted $\varnothing$, is the set with no member.

Definition 0.2. Let S be a given set, and $A \subseteq S$, $B \subseteq S$. The set $A \cup B$, called the ***union*** of A and B, consists of members belonging to set A or set B, and the set $A \cap B$, called the ***intersection*** of A and B, consists of members belonging to both set A and set B.

Let $\mathscr{F}$ be a collection of subsets in S. The set $\bigcup\limits_{A \in \mathscr{F}} A$, called the ***union*** of sets in $\mathscr{F}$, is defined by

$$\bigcup_{A \in \mathscr{F}} A = \{x \in S \,|\, (\exists\, A \in \mathscr{F})(x \in A)\},$$

and $\bigcap\limits_{A \in \mathscr{F}} A = \{x \in S \,|\, (\forall\, A \in \mathscr{F})(x \in A)\}$ is the ***intersection*** of sets in $\mathscr{F}$. When $\mathscr{F} = \{A_\alpha \,|\, \alpha \in I\}$, the union and the intersection of sets in $\mathscr{F}$ can also be written as $\bigcup\limits_{\alpha \in I} A_\alpha$ and

$\bigcap\limits_{\alpha \in I} A_\alpha$, respectively. When $\mathscr{F} = \big\{A_1, A_2, \cdots, A_N\big\}$, the union and the intersection of sets in $\mathscr{F}$ can also be written as $\bigcup\limits_{i=1}^{N} A_i$ and $\bigcap\limits_{i=1}^{N} A_i$, respectively.

Example 0.3. Let $\mathscr{F}$ be the collection of open intervals with length 2 and mid-point in $[0, 1]$. Then $\mathscr{F} = \big\{(\alpha - 1, \alpha + 1) \,\big|\, \alpha \in [0, 1]\big\}$. Moreover,

$$\bigcup_{A \in \mathscr{F}} A = \bigcup_{\alpha \in [0,1]} (\alpha - 1, \alpha + 1) = (-1, 2) \quad \text{and} \quad \bigcap_{A \in \mathscr{F}} A = \bigcap_{\alpha \in [0,1]} (\alpha - 1, \alpha + 1) = (0, 1).$$

Definition 0.4. Let S be a given set, and $A \subseteq S$, $B \subseteq S$. The ***complement*** of A relative to B, denoted $B \backslash A$, is the set consisting of members of B that are not members of A. When the universal set S under consideration is fixed, the complement of A relative to S or simply the complement of A, is denoted by $A^{\complement}$, or $S \backslash A$.

Theorem 0.5. (De Morgan's Law)

1. $B \backslash \bigcup\limits_{\alpha \in I} A_\alpha = \bigcap\limits_{\alpha \in I} (B \backslash A_\alpha)$ *or* $B \backslash \bigcup\limits_{A \in \mathscr{F}} A = \bigcap\limits_{A \in \mathscr{F}} (B \backslash A)$.
2. $B \backslash \bigcap\limits_{\alpha \in I} A_\alpha = \bigcup\limits_{\alpha \in I} (B \backslash A_\alpha)$ *or* $B \backslash \bigcap\limits_{A \in \mathscr{F}} A = \bigcup\limits_{A \in \mathscr{F}} (B \backslash A)$.

Proof. By definition,

$$\begin{aligned} x \in B \backslash \bigcup_{\alpha \in I} A_\alpha &\Leftrightarrow x \in B \text{ and } x \notin \bigcup_{\alpha \in I} A_\alpha \Leftrightarrow x \in B \text{ and } x \notin A_\alpha \text{ for all } \alpha \in I \\ &\Leftrightarrow x \in B \backslash A_\alpha \text{ for all } \alpha \in I \Leftrightarrow x \in \bigcap_{\alpha \in I} (B \backslash A_\alpha) \end{aligned}$$

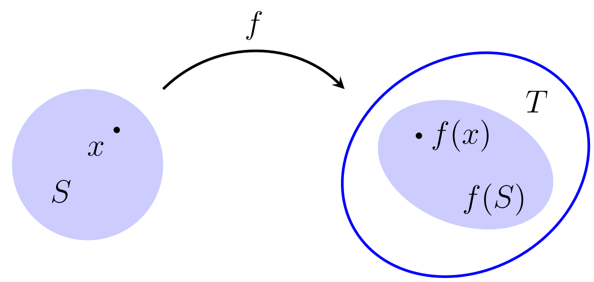
The proof of the second identity is similar, and is left as an exercise. □

Definition 0.6. An ***ordered pair*** (a, b) is an object formed from two objects a and b, where a is called the first coordinate and b the second coordinate. Two ordered pairs are equal whenever their corresponding coordinates are the same. An ***ordered n-tuples*** $(a_1, a_2, \cdots, a_n)$ is an object formed from n objects $a_1, a_2, \cdots, a_n$, where for each j, a_j is called the j-th coordinate. Two n-tuples $(a_1, a_2, \cdots, a_n)$, $(c_1, c_2, \cdots, c_n)$ are equal if $a_j = c_j$ for all $j \in \{1, \cdots, n\}$.

Definition 0.7. Given sets A and B, the ***Cartesian product*** $A \times B$ of A and B is the set of all ordered pairs (a, b) with $a \in A$ and $b \in B$, $A \times B = \big\{(a, b) \,\big|\, a \in A \text{ and } b \in B\big\}$. The Cartesian of three or more sets are defined similarly.

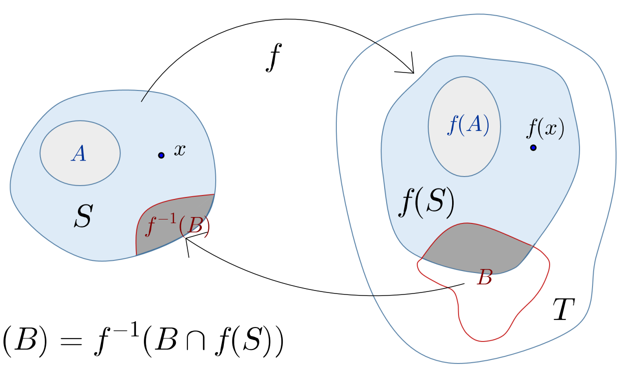
0.2 Functions

Definition 0.8. Let S and T be given sets. A ***function*** $f : S \to T$ consists of two sets S and T together with a "rule" that assigns to each $x \in S$ a special element of T denoted by $f(x)$. One writes $x \mapsto f(x)$ to denote that x is mapped to the element $f(x)$. S is called the ***domain*** (定義域) of f, and T is called the ***target*** or ***co-domain*** (對應域) of f. The ***range*** (值域) of f or the ***image*** of f, is the subset of T defined by $f(S) = \{f(x) \,|\, x \in S\}$.



Definition 0.9. A function $f : S \to T$ is called ***one-to-one*** (一對一), ***injective*** or an ***injection*** if $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$ (which is equivalent to that $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$). A function $f : S \to T$ is called ***onto*** (映成), ***surjective*** or an ***surjection*** if $\forall\, y \in T$, $\exists\, x \in S$, $\ni f(x) = y$ (that is, $f(S) = T$). A function $f : S \to T$ is called an ***bijection*** if it is one-to-one and onto.

Definition 0.10. For $f : S \to T$, $A \subseteq S$, we call $f(A) = \{f(x) \,|\, x \in A\}$ the ***image*** of A under f. For $B \subseteq T$, we call $f^{-1}(B) = \{x \in S \,|\, f(x) \in B\}$ the ***pre-image*** of B under f.



$$f^{-1}(B) = f^{-1}(B \cap f(S))$$

Proposition 0.11. *Let $f : S \to T$ be a function, C_1 ,$C_2 \subseteq T$ and D_1, $D_2 \subseteq S$.*

(a) $f^{-1}(C_1 \cup C_2) = f^{-1}(C_1) \cup f^{-1}(C_2)$.

(b) $f(D_1 \cup D_2) = f(D_1) \cup f(D_2)$.

(c) $f^{-1}(C_1 \cap C_2) = f^{-1}(C_1) \cap f^{-1}(C_2)$.

(d) $f(D_1 \cap D_2) \subseteq f(D_1) \cap f(D_2)$.

(e) $f^{-1}(f(D_1)) \supseteq D_1$ (*"=" if f is one-to-one*).

(f) $f(f^{-1}(C_1)) \subseteq C_1$ (*"=" if $C_1 \subseteq f(S)$*).

Proof. We only prove (c) and (d), and the proof of the other statements are left as an exercise.

(c) We first show that $f^{-1}(C_1 \cap C_2) \subseteq f^{-1}(C_1) \cap f^{-1}(C_2)$. Suppose that $x \in f^{-1}(C_1 \cap C_2)$. Then $f(x) \in C_1 \cap C_2$. Therefore, $f(x) \in C_1$ and $f(x) \in C_2$, or equivalently, $x \in f^{-1}(C_1)$ and $x \in f^{-1}(C_2)$; thus $x \in f^{-1}(C_1) \cap f^{-1}(C_2)$.

Next, we show that $f^{-1}(C_1) \cap f^{-1}(C_2) \subseteq f^{-1}(C_1 \cap C_2)$. Suppose that $x \in f^{-1}(C_1) \cap f^{-1}(C_2)$. Then $x \in f^{-1}(C_1)$ and $x \in f^{-1}(C_2)$ which implies that $f(x) \in C_1$ and $f(x) \in C_2$; thus $f(x) \in C_1 \cap C_2$ or equivalently, $x \in f^{-1}(C_1 \cap C_2)$.

(d) Suppose that $y \in f(D_1 \cap D_2)$. Then there exists $x \in D_1 \cap D_2$ such that $y = f(x)$. As a consequence, $y \in f(D_1)$ and $y \in f(D_2)$ which implies that $y \in f(D_1) \cap f(D_2)$. □

0.3 Countability of Sets

Definition 0.12. A set S is called ***denumerable*** or ***countably infinite*** (無窮可數的) if S can be put into one-to-one correspondence with $\mathbb{N}$; that is, S is denumerable if and only if there exists $f : \mathbb{N} \to S$ which is one-to-one and onto. A set is called ***countable*** (可數的) if it is either finite or denumerable, and is called ***uncountable*** if it is not countable.

Remark 0.13. If $f : \mathbb{N} \xrightarrow[onto]{1-1} S$, then $f^{-1} : S \xrightarrow[onto]{1-1} \mathbb{N}$. Therefore,

$$S \text{ is denumerable} \Leftrightarrow \exists\, f : \mathbb{N} \xrightarrow[onto]{1-1} S \Leftrightarrow \exists\, g = f^{-1} : S \xrightarrow[onto]{1-1} \mathbb{N}.$$

f can be thought as a rule of counting/labeling elements in S since $S = \{f(1), f(2), \cdots\}$.

Example 0.14. $\mathbb{N}$ is countable since $f : \mathbb{N} \xrightarrow[onto]{1-1} \mathbb{N}$ with $f(x) = x$, $\forall\, n \in \mathbb{N}$.

Example 0.15. $\mathbb{Z}$ is countable. $f : \mathbb{Z} \to \mathbb{N}$ with $f(x) = \begin{cases} 1 & \text{if } x = 0 \\ 2x & \text{if } x > 0 \\ -2x+1 & \text{if } x < 0 \end{cases}$.

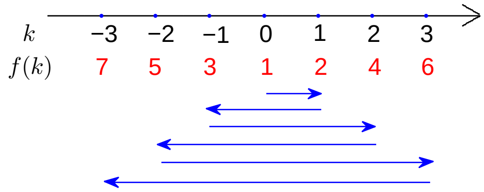


Figure 1: An illustration of how elements in $\mathbb{Z}$ are labeled

Example 0.16. The set $\mathbb{N} \times \mathbb{N} = \{(a, b) \,|\, a, b \in \mathbb{N}\}$ is countable. In fact, two ways of mapping are shown in the figures below.

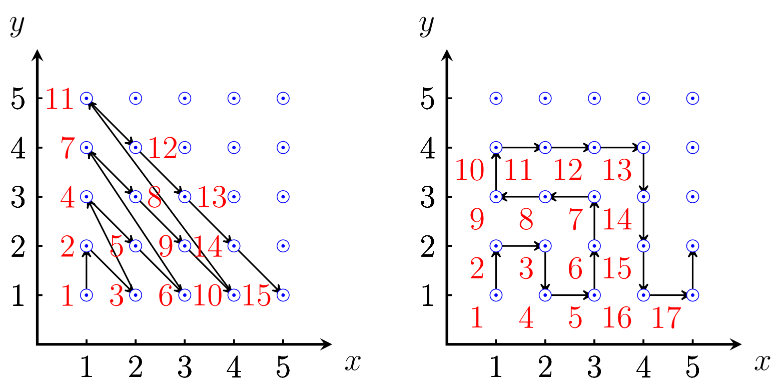


Figure 2: The illustration of two ways of labeling elements in $\mathbb{N} \times \mathbb{N}$

Proposition 0.17. *Let S be a non-empty set. The following three statements are equivalent:*

(a) *S is countable;*

(b) *there exists a surjection $f : \mathbb{N} \to S$;*

(c) *there exists an injection $f : S \to \mathbb{N}$.*

Proof. "(a) $\Rightarrow$ (b)" First suppose that $S = \{x_1, \cdots, x_n\}$ is finite. Define $f : \mathbb{N} \to S$ by

$$f(k) = \begin{cases} x_k & \text{if } k < n, \\ x_n & \text{if } k \geqslant n. \end{cases}$$

Then $f : \mathbb{N} \to S$ is a surjection. Now suppose that S is denumerable. Then by definition of countability, there exists $f : \mathbb{N} \xrightarrow[\text{onto}]{1-1} S$.

"(a) $\Leftarrow$ (b)" W.L.O.G. (without loss of generality, 不失一般性) we assume that S is an infinite set. Let $k_1 = 1$. Since $\#(S) = \infty$, $S_1 \equiv S \backslash \{f(k_1)\} \neq \varnothing$; thus $N_1 \equiv f^{-1}(S_1)$ is a non-empty subset of $\mathbb{N}$. By the well-ordered property of $\mathbb{N}$ (that is, non-empty subset of $\mathbb{N}$ has least element), N_1 has a smallest element denoted by k_2. Since $\#(S) = \infty$, $S_2 = S \backslash \{f(k_1), f(k_2)\} \neq \varnothing$; thus $N_2 \equiv f^{-1}(S_2)$ is a non-empty subset of $\mathbb{N}$ and possesses a smallest element denoted by k_3. We continue this process and obtain a set $\{k_1, k_2, \cdots\} \subseteq \mathbb{N}$, where $k_1 < k_2 < \cdots$, and k_j is the smallest element of $N_{j-1} \equiv f^{-1}(S \backslash \{f(k_1), f(k_2), \cdots, f(k_{j-1})\})$.

Claim: $f : \{k_1, k_2, \cdots\} \to S$ is one-to-one and onto.

Proof of claim: The injectivity of f is due to that $f(k_j) \notin \big\{f(k_1), f(k_2), \cdots, f(k_{j-1})\big\}$ for all $j \geqslant 2$. For surjectivity, assume that there is $s \in S$ such that $s \notin f(\{k_1, k_2, \cdots\})$. Since $f : \mathbb{N} \to \mathbb{S}$ is onto, $f^{-1}(\{s\})$ is a non-empty subset of $\mathbb{N}$; thus possesses a smallest element k. Since $s \notin f(\{k_1, k_2, \cdots\})$, there exists $\ell \in \mathbb{N}$ such that $k_\ell < k < k_{\ell+1}$. As a consequence, there exists $k \in N_\ell$ such that $k < k_{\ell+1}$ which contradicts to the fact that $k_{\ell+1}$ is the smallest element of N_ℓ.

Define $g : \mathbb{N} \to \{k_1, k_2, \cdots\}$ by $g(j) = k_j$. Then $g : \mathbb{N} \to \{k_1, k_2, \cdots\}$ is one-to-one and onto; thus $h = g \circ f : \mathbb{N} \xrightarrow[onto]{1-1} S$.

"(a) $\Rightarrow$ (c)" If $S = \{x_1, \cdots, x_n\}$ is finite, we simply let $f : S \to \mathbb{N}$ be $f(x_n) = n$. Then f is clearly an injection. If S is denumerable, by definition there exists $g : \mathbb{N} \xrightarrow[onto]{1-1} S$ which shows that $f = g^{-1} : S \to \mathbb{N}$ is an injection.

"(a) $\Leftarrow$ (c)" Let $f : S \to \mathbb{N}$ be an injection. If f is also surjective, then $f : S \xrightarrow[onto]{1-1} \mathbb{N}$ which implies that S is denumerable. Now suppose that $f(S) \subsetneq \mathbb{N}$. Since S is non-empty, there exists $s \in S$. Let $g : \mathbb{N} \to S$ be defined by

$$g(n) = \begin{cases} f^{-1}(n) & \text{if } n \in f(S), \\ s & \text{if } n \notin f(S). \end{cases}$$

Then clearly $g : \mathbb{N} \to S$ is surjective; thus the equivalence between (a) and (b) implies that S is countable. □

Example 0.18. The set $\mathbb{N} \times \mathbb{N}$ is countable since the map $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by $f\big((m, n)\big) = 2^m 3^n$ is an injection.

Theorem 0.19. *Any non-empty subset of a countable set is countable.*

Proof. Let S be a countable set, and A be a non-empty subset of S. Since S is countable, by by Proposition 0.17 there exists a surjection $f : \mathbb{N} \to S$. On the other hand, since A is a non-empty subset of S, there exists $a \in A$. Define

$$g(x) = \begin{cases} x & \text{if } x \in A, \\ a & \text{if } x \notin A. \end{cases}$$

Then $h = g \circ f : \mathbb{N} \to A$ is a surjection, and Proposition 0.17 implies that A is countable. ▫

Theorem 0.20. *The union of countable countable sets is countable.* (可數個可數集的聯集是可數的)

Proof. Let A_i be countable, and define $A = \bigcup\limits_{i=1}^{\infty} A_i$. Write $A_i = \{x_{i1}, x_{i2}, x_{i3}, \cdots\}$. Then $A = \big\{x_{ij} \,\big|\, i = 1, 2, \cdots,\ j < \#(A_i) + 1\big\}$, where $\#(A_i) = \infty$ if A_i is countably infinite. Let $S = \big\{(i,j) \,\big|\, i = 1, 2, \cdots, j < \#(A_i) + 1\big\}$, and define $f : S \to A$ by $f\big((i,j)\big) = x_{ij}$. Then $f : S \to A$ is a surjection. On the other hand, since S is a subset of $\mathbb{N} \times \mathbb{N}$, Theorem 0.19 implies that S is countable; thus Proposition 0.17 guarantees the existence of a surjection $g : \mathbb{N} \to S$. Then $h = f \circ g : \mathbb{N} \to A$ is a surjection which, by Proposition 0.17 again, implies that A is countable. ▫

Example 0.21. $\mathbb{Z} \times \mathbb{Z}$ is countable.

Proof. For $i \in \mathbb{Z}$, let $A_i = \big\{(i,j) \,\big|\, j \in \mathbb{Z}\big\}$. By Example 0.15, A_i is countable for all $i \in \mathbb{Z}$. Since $\mathbb{Z} \times \mathbb{Z} = \bigcup_{i \in \mathbb{Z}} A_i$ which is countable union of countable sets, Theorem 0.20 implies that $\mathbb{Z} \times \mathbb{Z}$ is countable. ▫

Theorem 0.22. $\mathbb{Q}$ *is countable.*

Proof. Define

$$f(x) = \begin{cases} (p,q), & \text{if } x > 0, \quad x = \dfrac{q}{p}, \quad \gcd(p,q) = 1,\ p > 0. \\ (0,0), & \text{if } x = 0. \\ (p,-q), & \text{if } x < 0, \quad x = -\dfrac{q}{p}, \quad \gcd(p,q) = 1,\ p > 0. \end{cases}$$

Then $f : \mathbb{Q} \to \mathbb{Z} \times \mathbb{Z}$ is one-to-one; thus $f : \mathbb{Q} \xrightarrow[onto]{1-1} f(\mathbb{Q})$. Since $\mathbb{Z} \times \mathbb{Z}$ is countable, its non-empty subset $f(\mathbb{Q})$ is also countable. As a consequence, there exists $g : f(\mathbb{Q}) \xrightarrow[onto]{1-1} \mathbb{N}$; thus $h = g \circ f : \mathbb{Q} \xrightarrow[onto]{1-1} \mathbb{N}$. ▫

Theorem 0.23. *The open interval* $(0,1)$ *is uncountable.*

Proof. Assume the contrary that there exists $f:\mathbb{N}\to(0,1)$ which is one-to-one and onto. Write $f(k)$ in decimal expansion (十進位展開); that is,

$$\begin{aligned} f(1) &= 0.d_{11}d_{21}d_{31}\cdots \\ f(2) &= 0.d_{12}d_{22}d_{32}\cdots \\ \vdots\ & \qquad\ \vdots \\ f(k) &= 0.d_{1k}d_{2k}d_{3k}\cdots \\ \vdots\ & \qquad\ \vdots \end{aligned}$$

Here we note that repeated 9's are chosen by preference over terminating decimals; that is, for example, we write $\dfrac{1}{4}=0.249999\cdots$ instead of $\dfrac{1}{4}=0.250000\cdots$.

Let $x\in(0,1)$ be such that $x=0.d_1d_2\cdots$, where

$$d_k=\begin{cases} 5 & \text{if} \quad d_{kk}\neq 5\,, \\ 3 & \text{if} \quad d_{kk}=5\,. \end{cases}$$

（建構一個 x 使其小數點下第 k 位數與 $f(k)$ 的小數點下第 k 位數不相等）. Then $x\neq f(k)$ for all $k\in\mathbb{N}$, a contradiction; thus $(0,1)$ is uncountable. □

Corollary 0.24. *The collection of real numbers is uncountable.*

0.4 Exercises

§ 0.1 Sets

Problem 0.1. Let A, B, C be given sets. Show that

1. $(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$.

2. $(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$.

§ 0.2 Functions

Problem 0.2. Let S and T be given sets, $A\subseteq S$, $B\subseteq T$, and $f:S\to T$. Show that

1. $f(f^{-1}(B))\subseteq B$, and $f(f^{-1}(B))=B$ if $B\subseteq f(S)$.

2. $f^{-1}(f(A)) \supseteq A$, and $f^{-1}(f(A)) = A$ if $f : S \to T$ is one-to-one.

Problem 0.3. Let A and B be two non-empty sets and $f : A \to B$. Show that

1. $f^{-1}(D_1 \cup D_2) = f^{-1}(D_1) \cup f^{-1}(D_2)$, $f^{-1}(D_1 \cap D_2) = f^{-1}(D_1) \cap f^{-1}(D_2)$, $f^{-1}(D_1 \backslash D_2) = f^{-1}(D_1) \backslash f^{-1}(D_2)$ for all $D_1, D_2 \subseteq B$.

2. $f(C_1 \cup C_2) = f(C_1) \cup f(C_2)$, $f(C_1 \cap C_2) \subseteq f(C_1) \cap f(C_2)$ for all $C_1, C_2 \subseteq A$.

Problem 0.4. Let A and B be two non-empty sets and $f : A \to B$. Show that the following three statements are equivalent; that is, show that each one of the following statements implies the other four.

1. f is one-to-one.

2. For every y in B, the set $f^{-1}(\{y\})$ contains at most one point.

3. $f^{-1}(f(C)) = C$ for all $C \subseteq A$.

4. $f(C_1 \cap C_2) = f(C_1) \cap f(C_2)$ for all subsets C_1 and C_2 of A.

5. $f(C_2 \backslash C_1) = f(C_2) - f(C_1)$ for all $C_1 \subseteq C_2 \subseteq A$.

§ 0.3 Countability of Sets

Problem 0.5 (True or False). Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. Given two sets A and B. Then $A \times B$ is countable if and only if A and B are countable.

2. The set $\{(x, y) \in \mathbb{R}^2 \,|\, x + y \in \mathbb{Q}\}$ is countable.

3. Let $A \subseteq \mathbb{R}$ satisfy

$$\sup \Big\{ \sum_{b \in B} |b| \,\Big|\, B \text{ is a non-empty finite subsets of } A \Big\} < \infty .$$

Then $\{x \in A \,|\, x \neq 0\}$ is countable.

Chapter 1

The Real Number System and Completeness

1.1 Ordered Fields

Definition 1.1. A set $\mathbb{F}$ is said to be a ***field*** (體) if there are two operations $+$ and $\cdot$ such that

1. $x + y \in \mathbb{F}$, $x \cdot y \in \mathbb{F}$ if $x, y \in \mathbb{F}$. (封閉性)

2. $x + y = y + x$ for all $x, y \in \mathbb{F}$. (commutativity, 加法的交換性)

3. $(x + y) + z = x + (y + z)$ for all $x, y, z \in \mathbb{F}$. (associativity, 加法的結合性)

4. There exists $0 \in \mathbb{F}$, called the additive identity (加法單位元素), such that $x + 0 = x$ for all $x \in \mathbb{F}$. (the existence of zero)

5. For every $x \in \mathbb{F}$, there exists $y \in \mathbb{F}$ (usually y is denoted by $-x$ and is called the additive inverse (加法反元素) of x) such that $x + y = 0$. One writes $x - y \equiv x + (-y)$.

6. $x \cdot y = y \cdot x$ for all $x, y \in \mathbb{F}$. (乘法的交換性)

7. $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ for all $x, y, z \in \mathbb{F}$. (乘法的結合性)

8. There exists $1 \in \mathbb{F}$, called the multiplicative identity (乘法單位元素), such that $x \cdot 1 = x$ for all $x \in \mathbb{F}$. (the existence of unity)

9. For every $x \in \mathbb{F}$, $x \neq 0$, there exists $y \in \mathbb{F}$ (usually y is denoted by x^{-1}, and is

called the multiplicative inverse (乘法反元素) of x) such that $x \cdot y = 1$. One writes $x \cdot y \equiv x \cdot x^{-1} = 1$.

10. $x \cdot (y + z) = x \cdot y + x \cdot z$ for all $x, y, z \in \mathbb{F}$. (distributive law, 分配律)
11. $0 \neq 1$.

Remark 1.2. Let x and y be both multiplicative inverse（乘法反元素）of a number a in $(\mathbb{F}, +, \cdot)$. Then

$$x \cdot a = 1 \quad \Rightarrow \quad (x \cdot a) \cdot y = 1 \cdot y = y \quad \Rightarrow \quad x \cdot 1 = x \cdot (a \cdot y) = y;$$

thus $x = y$. In other words, the multiplicative inverse of a number is unique. Similarly, the additive inverse of a number is also unique.

Remark 1.3. A set $\mathbb{F}$ satisfying properties 1 to 10 with $0 = 1$ consists of only one member: By distributive law, $x \cdot 0 = x \cdot (0+0) = x \cdot 0 + x \cdot 0$; thus $-(x \cdot 0) + (x \cdot 0) = -(x \cdot 0) + (x \cdot 0) + (x \cdot 0)$ which implies that $x \cdot 0 = 0$. Therefore, if $0 = 1$, then $x = x \cdot 1 = x \cdot 0 = 0$ for all $x \in \mathbb{F}$. Hence, the set $\mathbb{F}$ consists only one element 0.

Remark 1.4. If $x \in \mathbb{F}$, then $\big((1 + (-1)\big) \cdot x = 0$ which implies that $x + (-1) \cdot x = 0$. Therefore, $(-1) \cdot x = -x + x + (-1) \cdot x = -x + 0 = -x$.

Example 1.5. Let $\mathbb{F} = \{a, b, c\}$ with the operations $+$ and $\cdot$ defined by

$$\begin{array}{c|ccc} + & a & b & c \\ \hline a & a & b & c \\ b & b & c & a \\ c & c & a & b \end{array} \qquad \begin{array}{c|ccc} \cdot & a & b & c \\ \hline a & a & a & a \\ b & a & b & c \\ c & a & c & b \end{array}.$$

Then $\mathbb{F}$ is a field because of the following: Properties 1, 2, 3, 6, 7 are obvious.
Property 4: $\exists$ "0" $\ni x +$ "0" $= x$ for all $x \in \mathbb{F}$. In fact, "0" $= a$.
Property 5: $\forall\, x \in \mathbb{F}$, $\exists\, y \in \mathbb{F} \ni x + y = 0$, here $b = -c$, $c = -b$.
Property 8: $\exists$ "1" $\ni x \cdot$ "1" $= x$ for all $x \in \mathbb{F}$. In fact, "1" $= b$ (so Property 11 holds since $a \neq b$).
Property 9: $\forall\, x \neq 0, \in \mathbb{F}$, $\exists\, z \in \mathbb{F} \ni x \cdot z = 1$, here $z = x$.
The validity of Property 10 is left as an exercise.

Example 1.6. Let $(\mathbb{F}, +, \cdot)$ be a field. Consider the set $\mathcal{F} = \mathbb{F} \times \mathbb{F} = \big\{(a, b) \,\big|\, a, b \in \mathbb{F}\big\}$. Define

$$(a, b) \oplus (c, d) = (a + c, b + d) \quad \text{and} \quad (a, b) \odot (c, d) = (a \cdot c - b \cdot d, a \cdot d + b \cdot c).$$

Then $(\mathcal{F}, \oplus, \odot)$ is also a field. The ordered pair (a, b) in $\mathcal{F}$ is sometimes denoted by $a + bi$.

Example 1.7. Let $(\mathbb{F}, +, \cdot)$ be a field. Then $(x-y)(x+y) = x^2 - y^2$ for all $x, y \in \mathbb{F}$. In fact,

$$\begin{aligned}
(x-y)(x+y) &= (x-y)\cdot x + (x-y)\cdot y && \text{(by 分配律)}\\
&= x\cdot(x-y) + y\cdot(x-y) && \text{(by 乘法交換律)}\\
&= x\cdot x + x\cdot(-y) + y\cdot x + y\cdot(-y) && \text{(by 分配律)}\\
&= x^2 - x\cdot y + x\cdot y - y^2 && \text{(by Remark 1.4 and 乘法交換律)}\\
&= x^2 + 0 - y^2 && \text{(by Property 5)}\\
&= x^2 - y^2 && \text{(by Property 4).}
\end{aligned}$$

Definition 1.8. An ***ordered field*** (有序體) is a field $(\mathbb{F}, +, \cdot)$ equipped with a relation $\leqslant$ on $\mathbb{F}$ satisfying that

1. $x \leqslant x$ for all $x \in \mathbb{F}$ (reflexivity).
2. If $x, y \in \mathbb{F}$ satisfies that $x \leqslant y$ and $y \leqslant x$, then $x = y$ (anti-symmetry).
3. If $x, y \in \mathbb{F}$ satisfies that $x \leqslant y$ and $y \leqslant z$, then $x \leqslant z$ (transitivity).
4. For each $x, y \in \mathbb{F}$, either $x \leqslant y$ or $y \leqslant x$.
5. If $x \leqslant y$, then $x + z \leqslant y + z$ for all $z \in \mathbb{F}$ (compatibility of $\leqslant$ and $+$).
6. If $0 \leqslant x$ and $0 \leqslant y$, then $0 \leqslant x \cdot y$ (compatibility of $\leqslant$ and $\cdot$).

Remark 1.9. A relation $\leqslant$ on a field $\mathbb{F}$ satisfying only 1-3 in the definition above is called a partial order. If in addition $\leqslant$ also satisfies 4, it is called a total order or linear order. Note that $\geqslant$ is a relation on $\mathbb{Q}$ satisfying 1-5 but not 6.

Remark 1.10. In an ordered field, the multiplicative inverse of $x \neq 0$ is sometimes denoted by $\dfrac{1}{x}$.

Definition 1.11. In an ordered field $(\mathbb{F}, +, \cdot, \leqslant)$, the binary relations $<$, $\geqslant$ and $<$ are defined by:

1. $x < y$ if $x \leqslant y$ and $x \neq y$.
2. $x \geqslant y$ if $y \leqslant x$.
3. $x > y$ if $y < x$.

From now on, the total order $\preccurlyeq$ of an ordered field will be denoted by $\leqslant$, and the symbols $\prec$, $\succcurlyeq$ and $\succ$ will be denoted by $<$, $\geqslant$ and $>$, respectively.

Adopting the definition above, it is not immediately clear that $x \nleqslant y \Leftrightarrow x > y$. However, this is indeed the case, and to be more precise we have the following

Proposition 1.12. (Law of Trichotomy, 三一律) *If x and y are elements of an ordered field $(\mathbb{F}, +, \cdot, \leqslant)$, then exactly one of the relations $x < y$, $x = y$ or $y < x$ holds.*

Proof. Since $\mathbb{F}$ is a totally ordered field, x and y are comparable. Therefore, either $x \leqslant y$ or $y \leqslant x$. Assume that $x \leqslant y$.

1. If $x = y$, then $x \nless y$ and $x \ngtr y$.

2. If $x \neq y$, then $x < y$. If it also holds that $x > y$, then $x \geqslant y$; thus by the property of anit-symmetry of an order, we must have $x = y$, a contradiction. Therefore, it can only be that $x < y$.

The proof for the case $y \leqslant x$ is similar, and is left as an exercise. ◻

Proposition 1.13. *Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, and $a, b, x, y, z \in \mathbb{F}$.*

1. *If $a + x = a$, then $x = 0$.*
   If $a \cdot x = a$ and $a \neq 0$, then $x = 1$.

2. *If $a + x = 0$, then $x = -a$.*
   If $a \cdot x = 1$ and $a \neq 0$, then $x = a^{-1}$.

3. *If $x \cdot y = 0$, then $x = 0$ or $y = 0$.*

4. *If $x \leqslant y < z$ or $x < y \leqslant z$, then $x < z$ (the transitivity of $<$).*

5. *If $a < b$, then $a + x < b + x$ (the compatibility of $<$ and $+$).*
   If $0 < a$ and $0 < b$, then $0 < a \cdot b$ (the compatibility of $<$ and $\cdot$).

6. *If $a + x = b + x$, then $a = b$.*
   If $a + x \leqslant (<)\, b + x$, then $a \leqslant (<)\, b$.
   If $a \cdot x = b \cdot x$ and $x \neq 0$, then $a = b$.
   If $a \cdot x \leqslant (<)\, b \cdot x$ and $x > 0$, then $a \leqslant (<)\, b$.

7. $0 \cdot x = 0$.

8. $-(-x) = x$.

9. $-x = (-1) \cdot x$.

10. *If* $x \neq 0$, *then* $x^{-1} \neq 0$ *and* $(x^{-1})^{-1} = x$.

11. *If* $x \neq 0$ *and* $y \neq 0$, *then* $x \cdot y \neq 0$ *and* $(x \cdot y)^{-1} = x^{-1} \cdot y^{-1}$.

12. *If* $x \leqslant (<) \, y$ *and* $0 \leqslant (<) \, z$, *then* $x \cdot z \leqslant (<) \, y \cdot z$.
    If $x \leqslant (<) \, y$ *and* $0 \geqslant (>) \, z$, *then* $x \cdot z \geqslant (>) \, y \cdot z$.

13. *If* $x \leqslant (<) \, 0$ *and* $y \leqslant (<) \, 0$, *then* $x \cdot y \geqslant (>) \, 0$.
    If $x \leqslant (<) \, 0$ *and* $y \geqslant (>) \, 0$, *then* $x \cdot y \leqslant (<) \, 0$.

14. $0 < 1$ *and* $-1 < 0$.

15. $x \cdot x \equiv x^2 \geqslant 0$.

16. *If* $x > 0$, *then* $x^{-1} > 0$. *If* $x < 0$, *then* $x^{-1} < 0$.

Proof. 1. $(-a) + a + x = (-a) + a = 0 \Rightarrow x = 0$.
$(a^{-1}) \cdot a \cdot x = (a^{-1}) \cdot a = 1 \Rightarrow x = 1$.

2. $(-a) + a + x = (-a) + 0 = -a \Rightarrow x = -a$.
   $(a^{-1}) \cdot a \cdot x = (a^{-1}) \cdot 1 = a^{-1} \Rightarrow x = a^{-1}$.

3. Assume that $x \neq 0$, then $x^{-1} \cdot x \cdot y = x^{-1} \cdot 0 = 0 \Rightarrow y = 0$.
   Assume that $y \neq 0$, then $x \cdot y \cdot y^{-1} = 0 \cdot y^{-1} = 0 \Rightarrow x = 0$.

4 and 5 are Left as an exercise.

6. $a + 0 = a + x + (-x) = b + x + (-x) = b + 0 \Rightarrow a = b$.
   $a + 0 = a + x + (-x) \leqslant b + x + (-x) = b + 0 \Rightarrow a \leqslant b$ (compatibility of $\leqslant$ and $+$).
   $a \cdot x \cdot x^{-1} = b \cdot x \cdot x^{-1} \Rightarrow a = b$.
   Suppose the contrary that $b < a$. Then $0 = b + (-b) \leqslant a + (-b)$. Since $x > 0$, $x \geqslant 0$; thus
   $$0 \leqslant \big(a + (-b)\big) \cdot x = a \cdot x + (-b) \cdot x \,.$$
   As a consequence, $b \cdot x = 0 + b \cdot x \leqslant a \cdot x + (-b) \cdot x + b \cdot x = a \cdot x$. By assumption, we must have $a \cdot x = b \cdot x$ or $(a - b) \cdot x = 0$. Using 3, $x = 0$ (since $a \neq b$), a contradiction.

7. See Remark 1.3.

8. $(-x) + (-(-x)) = 0 = (-x) + x \Rightarrow x = -(-x)$.

9. See Remark 1.4.

10. Assume $x^{-1} = 0$, $1 = x \cdot x^{-1} = x \cdot 0 = 0$, a contradiction. Therefore, $x^{-1} \neq 0$; thus $(x^{-1})^{-1} \cdot x^{-1} = 1 = x \cdot x^{-1} \Rightarrow (x^{-1})^{-1} = x$ (by 4).

11. That $x \cdot y = 0$ cannot be true since it is against Property 3, so $x \cdot y \neq 0$. Moreover,
$$(x \cdot y)^{-1}(x \cdot y) = 1 = 1 \cdot 1 = (x \cdot x^{-1}) \cdot (y \cdot y^{-1}) = (x^{-1} \cdot y^{-1}) \cdot (x \cdot y);$$
thus $(x \cdot y)^{-1} = x^{-1} \cdot y^{-1}$ (by 4).

12. If $x \leqslant (<) \, y$, then $0 = x + (-x) \leqslant (<) \, y + (-x)$. Since $0 \leqslant (<) \, z$, by the compatibility of $\leqslant (<)$ and $\cdot$ we must have $0 \leqslant (<) \, (y + (-x)) \cdot z = y \cdot z + (-x) \cdot z$. Therefore, by the compatibility of $\leqslant (<)$ and $+$, $x \cdot z = 0 + x \cdot z \leqslant (<) \, y \cdot z + (-x) \cdot z + x \cdot z = y \cdot z$. The second statement can be proved in a similar fashion.

13. Left as an exercise.

14. If $1 \leqslant 0$, then compatibility of $\leqslant$ and $+$ implies that $0 \leqslant -1$. By the compatibility of $\leqslant$ and $\cdot$, using 8 and 9 we find that $0 \leqslant (-1) \cdot (-1) = -(-1) = 1$; thus we conclude that $1 = 0$, a contradiction. As a consequence, $0 < 1$; thus the compatibility of $<$ and $+$ implies that $-1 < 0$.

15. Left as an exercise.

16. If $x > 0$ but $x^{-1} \leqslant 0$, then $1 = x \cdot x^{-1} \leqslant x \cdot 0 = 0$, a contradiction. □

Proposition 1.14. *Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, and $x, y \in \mathbb{F}$.*

1. *If $0 \leqslant x < y$, then $x^2 < y^2$.*

2. *If $0 \leqslant x, y$ and $x^2 < y^2$, then $x < y$.*

Proof. 1. By definition of "$<$", $0 \leqslant x \leqslant y$ and $x \neq y$. Using 12 of Proposition 1.13,
$$x^2 \leqslant y \cdot x < y \cdot y = y^2 .$$
By the transitivity of $<$, we conclude that $x^2 < y^2$.

2. Note that $x \neq y$, for if not, then $x^2 - y^2 = 0$ which contradicts to the assumption $x^2 < y^2$. Assume that $y < x$, then 1 implies that $y^2 < x^2$, a contradiction. □

Remark 1.15. Proposition 1.14 can be summarized as follows: if $x, y \geqslant 0$, then

$$x < y \Leftrightarrow x^2 < y^2 .$$

Moreover, Example 1.7, Proposition 1.13 and Proposition 1.14 together imply that if $x, y \geqslant 0$, then $x \leqslant y$ if and only if $x^2 \leqslant y^2$.

Definition 1.16. The ***magnitude*** or the ***absolute value*** of x, denoted $|x|$, is defined as

$$|x| = \begin{cases} x & \text{if } x \geqslant 0\,, \\ -x & \text{if } x < 0\,. \end{cases}$$

Proposition 1.17. *Let* $(\mathbb{F}, +, \cdot, \leqslant)$ *be an ordered field. Then*

1. $|x| \geqslant 0$ *for all* $x \in \mathbb{F}$.

2. $|x| = 0$ *if and only if* $x = 0$.

3. $-|x| \leqslant x \leqslant |x|$ *for all* $x \in \mathbb{F}$.

4. $|x \cdot y| = |x| \cdot |y|$ *for all* $x, y \in \mathbb{F}$.

5. $|x + y| \leqslant |x| + |y|$ *for all* $x, y \in \mathbb{F}$ (***triangle inequality***, 三角不等式).

6. $\big||x| - |y|\big| \leqslant |x - y|$ *for all* $x, y \in \mathbb{F}$.

Proof. Left as an exercise. □

Definition 1.18. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field. The ***natural number system***, denoted by $\mathbb{N}$, is the collection of all the numbers 1, $1+1$, $1+1+1$, $1+1+\cdots+1$ and etc. in $\mathbb{F}$. We write $2 \equiv 1+1$, $3 \equiv 2+1$, and $n \equiv \underbrace{1+1+\cdots+1}_{\text{(n times)}}$. In other words, $\mathbb{N} = \{1, 2, 3, \cdots\}$. The ***integer number system***, denoted by $\mathbb{Z}$, is the set $\mathbb{Z} = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$. The ***rational number system***, denoted by $\mathbb{Q}$, is the collection of all numbers of the form $\dfrac{q}{p} \equiv q \cdot p^{-1}$ with $p, q \in \mathbb{Z}$ and $p \neq 0$; that is,

$$\mathbb{Q} = \Big\{x \in \mathbb{F} \,\Big|\, x = \frac{q}{p}, p, q \in \mathbb{Z}, p \neq 0\Big\}.$$

Theorem 1.19. *Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field. Then $\mathbb{Q}$ is an order field.*

- **Peano's Axiom for natural numbers**:

1. 1 is a natural number.
2. Every natural number has a unique successor which is a natural number (+1 is defined on natural numbers).
3. No two natural numbers have the same successor ($n + 1 = m + 1$ implies $n = m$).
4. 1 is not a successor for any natural number (1 is the "smallest" natural number).
5. If a property is possessed by 1 and is possessed by the successor of every natural number that possesses it, then the property is possessed by all natural numbers. (如果某個被自然數 1 所擁有的性質，也被其它擁有這個性質的自然數的下一個自然數所擁有，那麼所有的自然數都會擁有這個性質)

- **Principle of Mathematical Induction (PMI)**: If $S \subseteq \mathbb{N}$ has the property that

(1) $1 \in S$, and (2) $n + 1 \in S$ whenever $n \in S$,

then $S = \mathbb{N}$.

- **Principle of Complete Induction (PCI)**: If $S \subseteq \mathbb{N}$ has the property that

$$\forall\, n \in \mathbb{N},\ n \in S \text{ whenever } \{1, 2, \cdots, n - 1\} \subseteq S,$$

then $S = \mathbb{N}$.

- **Well-Ordering Principle (WOP)**: Every nonempty subset of $\mathbb{N}$ has a smallest element.

Theorem 1.20. **PMI**, **PCI** *and* **WOP** *are equivalent.*

Definition 1.21. An order field $(\mathbb{F}, +, \cdot, \leqslant)$ is said to satisfy ***Archimedean Property*** (**AP**) if for all $x \in \mathbb{F}$ there exists $n \in \mathbb{Z}$ such that $x < n$.

Example 1.22. The rational number system $\mathbb{Q}$ satisfies Archimedean Property. To see this, let $x \in \mathbb{Q}$ be given. If $x \leqslant 0$, we take $n = 1$. Otherwise if $0 < x = \frac{q}{p}$ with $p, q \in \mathbb{N}$, we take $n = q + 1$ and it is obvious that $\frac{q}{p} \leqslant q < q + 1 = n$.

1.2 Sequences in Ordered Fields

Definition 1.23. A ***sequence*** in a set S is a function $f : \mathbb{N} \to S$ (not necessary one-to-one or onto). The values of f are called the ***terms*** of the sequence, and $f(n)$ is called the n-th terms of the sequence.

Remark 1.24. A sequence in S is a countable list of elements in S arranged in a particular order, and is usually denoted by $\{f(n)\}_{n=1}^{\infty}$ or $\{x_n\}_{n=1}^{\infty}$ with $x_n = f(n)$.

Definition 1.25. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field. An ***"open" interval*** in $\mathbb{F}$ is a set of the form (a, b) which consists of all $x \in \mathbb{F}$ satisfying $a < x < b$. A ***"closed" interval*** in $\mathbb{F}$ is a set of the form $[a, b]$ which consists of all $x \in \mathbb{F}$ satisfying $a \leqslant x \leqslant b$.

Definition 1.26. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field. A sequence $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{F}$ is said to be ***convergent*** if there exists $x \in \mathbb{F}$ such that for every $\varepsilon > 0$ (and $\varepsilon \in \mathbb{F}$),

$$\#\big\{n \in \mathbb{N} \,\big|\, x_n \notin (x - \varepsilon, x + \varepsilon)\big\} < \infty .$$

Such an x is called a ***limit*** of the sequence. In logic notation,

$$\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{F} \text{ is convergent} \quad \Leftrightarrow \quad (\exists\, x \in \mathbb{F})(\forall\, \varepsilon > 0)\big(\#\{n \in \mathbb{N} \,|\, x_n \notin (x - \varepsilon, x + \varepsilon)\} < \infty\big).$$

If x is a limit of $\{x_n\}_{n=1}^{\infty}$, we say $\{x_n\}_{n=1}^{\infty}$ converges to x and write $x_n \to x$ as $n \to \infty$. If no such x exists we say that $\{x_n\}_{n=1}^{\infty}$ ***diverges*** (or the limit of $\{x_n\}_{n=1}^{\infty}$ does not exist).

Remark 1.27. The number N may depend on ε, and smaller ε usually requires larger N.

In the definition above, it could happen that there are two different limits of a convergent sequence. In fact, this is never the case because of the following

Proposition 1.28. *If $\{x_n\}_{n=1}^{\infty}$ is a sequence in an ordered field $\mathbb{F}$, and $x_n \to x$ and $x_n \to y$ as $n \to \infty$, then $x = y$.* (*The uniqueness of the limit*).

Proof. Assume the contrary that $x \neq y$. W.L.O.G. we may assume that $x < y$, and let $\varepsilon = \dfrac{y - x}{2} > 0$. Define

$$A_1 = \big\{n \in \mathbb{N} \,\big|\, x_n \notin (x - \varepsilon, x + \varepsilon)\big\} \quad \text{and} \quad A_2 = \big\{n \in \mathbb{N} \,\big|\, x_n \notin (y - \varepsilon, y + \varepsilon)\big\}.$$

Then by the definition of the convergence of sequences, $\#A_1 < \infty$ and $\#A_2 < \infty$. Let $N_1 = \max A_1$, $N_2 = \max A_2$ and $N = \max\{N_1, N_2\}$. Since A_1, A_2 are finite, $N < \infty$. On the other hand, $N + 1 \notin A_1 \cup A_2$ which implies that $x_{N+1} \in (x - \varepsilon, x + \varepsilon) \cap (y - \varepsilon, y + \varepsilon) = \varnothing$, a contradiction. $\square$

Notation: Since the limit of a convergent sequence is unique, if $\{x_n\}_{n=1}^\infty$ is a convergent sequence, we use $\lim\limits_{n\to\infty} x_n$, where n is a dummy index and can be change to other letters, to denote the limit of $\{x_n\}_{n=1}^\infty$.

Example 1.29. A ***permutation*** of a non-empty set A is a one-to-one function from A onto A. Let $\pi : \mathbb{N} \to \mathbb{N}$ be a permutation of $\mathbb{N}$, and $\{x_n\}_{n=1}^\infty$ be a convergent sequence in an ordered field $\mathbb{F}$. Then $\{x_{\pi(n)}\}_{n=1}^\infty$ is also convergent since if x is the limit of $\{x_n\}_{n=1}^\infty$ and $\varepsilon > 0$,

$$\#\{n \in \mathbb{N} \,|\, x_{\pi(n)} \notin (x-\varepsilon, x+\varepsilon)\} = \#\{n \in \mathbb{N} \,|\, x_n \notin (x-\varepsilon, x+\varepsilon)\} < \infty .$$

Proposition 1.30. *Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, $\{x_n\}_{n=1}^\infty \subseteq \mathbb{F}$ be a sequence, and $x \in \mathbb{F}$. Then $\lim\limits_{n\to\infty} x_n = x$ if and only if for every $\varepsilon > 0$, there exists $N > 0$ such that $|x_n - x| < \varepsilon$ whenever $n \geqslant N$. In logic notation,*

$$\lim_{n\to\infty} x_n = x \quad \Leftrightarrow \quad (\forall\, \varepsilon > 0)(\exists\, N > 0)(n \geqslant N \Rightarrow |x_n - x| < \varepsilon) .$$

Proof. "$\Rightarrow$" Let $\varepsilon > 0$ be given. Since $\lim\limits_{n\to\infty} x_n = x$, $\#\{n \in \mathbb{N} \,|\, x_n \notin (x-\varepsilon, x+\varepsilon)\} < \infty$. If $\#\{n \in \mathbb{N} \,|\, x_n \notin (x-\varepsilon, x+\varepsilon)\} > 0$, define $N = \max\{n \in \mathbb{N} \,|\, x_n \notin (x-\varepsilon, x+\varepsilon)\} + 1$, otherwise define $N = 1$. Then if $n \geqslant N$, $x_n \in (x-\varepsilon, x+\varepsilon)$ or equivalently,

$$|x_n - x| < \varepsilon \quad \text{whenever} \quad n \geqslant N .$$

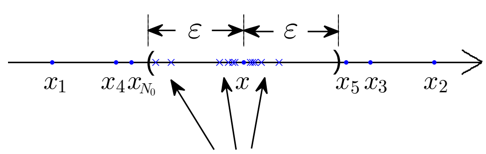


Figure 1.1: Let N_0 be the largest index of those x_n's outside $(x-\varepsilon, x+\varepsilon)$. Then $x_n \in (x-\varepsilon, x+\varepsilon)$ whenever $n \geqslant N = N_0 + 1$.

"$\Leftarrow$" Let $\varepsilon > 0$ be given. Then for some $N > 0$, if $n \geqslant N$, we have $|x_n - x| < \varepsilon$ or equivalently, if $n \geqslant N$, $x_n \in (x-\varepsilon, x+\varepsilon)$. This implies that

$$\#\{n \in \mathbb{N} \,|\, x_n \notin (x-\varepsilon, x+\varepsilon)\} < N < \infty . \qquad \square$$

Remark 1.31. By the proposition above, $x_n \to x$ as $n \to \infty$ if and only if the sequence $\{|x_n - x|\}_{n=1}^\infty$ converges to 0; that is,

$$\lim_{n\to\infty} x_n = x \qquad \text{if and only if} \qquad \lim_{n\to\infty} |x_n - x| = 0 .$$

Remark 1.32. A sequence $\{x_n\}_{n=1}^\infty \subseteq \mathbb{F}$ diverges if (and only if)

$$\forall\, x \in \mathbb{F}, \exists\, \varepsilon > 0 \ni \#\{n \in \mathbb{N} \,|\, x_n \notin (x-\varepsilon, x+\varepsilon)\} = \infty$$

which is equivalent to that

$$\forall\, x \in \mathbb{F}, \exists\, \varepsilon > 0 \ni \{n \in \mathbb{N} \,|\, x_n \notin (x-\varepsilon, x+\varepsilon)\} = \{n_1 < n_2 < \cdots < n_j < \cdots\}.$$

Therefore, $\{x_n\}_{n=1}^\infty$ diverges if (and only if)

$$\forall\, x \in \mathbb{F}, \exists\, \varepsilon > 0 \ni \forall\, N > 0, \exists\, n \geqslant N \text{ such that } |x_n - x| \geqslant \varepsilon.$$

Example 1.33. Now we use the ε-N argument as the definition of the convergence of sequences to re-establish the convergence of the sequence in Example 1.29.

Suppose that $\{x_n\}_{n=1}^\infty$ is a convergent sequence with limit x, and $\varepsilon > 0$ be given. Then there exists $N_1 > 0$ such that if $n \geqslant N_1$, we have $|x_n - x| < \varepsilon$. Let

$$N = \max\left\{\pi^{-1}(1), \pi^{-1}(2), \cdots, \pi^{-1}(N_1)\right\} + 1.$$

Then if $n \geqslant N$, $\pi(n) \geqslant N_1$ which implies that

$$\left|x_{\pi(n)} - x\right| < \varepsilon \quad \text{whenever} \quad n \geqslant N.$$

Therefore, $\lim\limits_{n\to\infty} x_{\pi(n)} = x$.

From the example above, we notice that proving the convergence using the ε-N argument seems more complicated; however, it is a necessary evil so we encourage the readers to major it.

Lemma 1.34 (Sandwich)**.** *If* $\lim\limits_{n\to\infty} x_n = L$, $\lim\limits_{n\to\infty} y_n = L$, $\{z_n\}_{n=1}^\infty$ *is a sequence such that* $x_n \leqslant z_n \leqslant y_n$, *then* $\lim\limits_{n\to\infty} z_n = L$.

Proof. Let $\varepsilon > 0$ be given. Since $\lim\limits_{n\to\infty} x_n = L$ and $\lim\limits_{n\to\infty} y_n = L$, by definition

$$\exists\, N_1 > 0 \ni L - \varepsilon < x_n < L + \varepsilon \ \text{ whenever } \ n \geqslant N_1$$

and

$$\exists\, N_2 > 0 \ni L - \varepsilon < y_n < L + \varepsilon \ \text{ whenever } \ n \geqslant N_2.$$

Let $N = \max\{N_1, N_2\}$. Then for $n \geqslant N$, $L - \varepsilon < x_n \leqslant z_n \leqslant y_n < L + \varepsilon$; thus $\lim\limits_{n\to\infty} z_n = L$. □

Proposition 1.35. *If* $x_n \leqslant y_n$ *for all* $n \in \mathbb{N}$ *and* $\lim\limits_{n\to\infty} x_n = x$, $\lim\limits_{n\to\infty} y_n = y$, *then* $x \leqslant y$.

Proof. Assume the contrary that $x > y$. Let $\varepsilon = \dfrac{x-y}{2}$. By the fact that $x_n \to x$ and $y_n \to y$ as $n \to \infty$, there exists $N_1, N_2 > 0$ such that

$$|x_n - x| < \frac{x-y}{2} \quad \text{whenever} \quad n \geqslant N_1 \quad \text{and} \quad |y_n - y| < \frac{x-y}{2} \quad \text{whenever} \quad n \geqslant N_2 .$$

Let $N = \max\{N_1, N_2\}$. Then for $n \geqslant N$,

$$y_n < y + \frac{x-y}{2} = \frac{x+y}{2} = x - \frac{x-y}{2} < x_n ,$$

a contradiction. □

Corollary 1.36. 1. *If $a \leqslant x_n$ (or $x_n \leqslant b$) and $\lim\limits_{n\to\infty} x_n = x$, then $a \leqslant x$ (or $x \leqslant b$).*

2. *If $a < x_n$ (or $x_n < b$) and $\lim\limits_{n\to\infty} x_n = x$, then $a \leqslant x$ (or $x \leqslant b$).*

Definition 1.37. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, and $\{x_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{F}$.

1. $\{x_n\}_{n=1}^{\infty}$ is said to be ***bounded from above***（有上界）if there exists $B \in \mathbb{F}$, called an ***upper bound*** of the sequence, such that $x_n \leqslant B$ for all $n \in \mathbb{N}$.

2. $\{x_n\}_{n=1}^{\infty}$ is said to be ***bounded from below***（有下界）if there exists $A \in \mathbb{F}$, called a ***lower bound*** of the sequence, such that $A \leqslant x_n$ for all $n \in \mathbb{N}$.

3. $\{x_n\}_{n=1}^{\infty}$ is said to be ***bounded***（有界的）if it is bounded from above and from below.

Remark 1.38. An equivalent definition of bounded sequences is stated as follows: $\{x_n\}_{n=1}^{\infty}$ is said to be bounded if there exists $M > 0$ such that $|x_n| \leqslant M$ for all $n \in \mathbb{N}$.

Proposition 1.39. *A convergent sequence is bounded*（數列收斂必有界）.

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a convergent sequence with limit x. Then there exists $N > 0$ such that

$$x_n \in (x-1, x+1) \qquad \forall\, n \geqslant N .$$

Let $M = \max\big\{|x_1|, |x_2|, \cdots, |x_{N-1}|, |x| + 1\big\}$. Then $|x_n| \leqslant M$ for all $n \in \mathbb{N}$. □

Theorem 1.40. *Suppose that $x_n \to x$ and $y_n \to y$ as $n \to \infty$. Then*

1. *$x_n \pm y_n \to x \pm y$ as $n \to \infty$.*

2. *$x_n \cdot y_n \to x \cdot y$ as $n \to \infty$.*

3. *If $y_n, y \neq 0$, then $\dfrac{x_n}{y_n} \to \dfrac{x}{y}$ as $n \to \infty$.*

Proof. 1. Let $\varepsilon > 0$ be given. Since $x_n \to x$ and $y_n \to y$ as $n \to \infty$, there exist $N_1, N_2 \in \mathbb{N}$ such that $|x_n - x| < \dfrac{\varepsilon}{2}$ for all $n \geqslant N_1$ and $|y_n - x| < \dfrac{\varepsilon}{2}$ whenever $n \geqslant N_2$. Define $N = \max\{N_1, N_2\}$. Then $N \in \mathbb{N}$ and if $n \geqslant N$,

$$|(x_n \pm y_n) - (x \pm y)| \leqslant |x_n - x\,| + |y_n - y\,| < \varepsilon;$$

thus $x_n \pm y_n \to x \pm y$ as $n \to \infty$.

2. Since $x_n \to x$ and $y_n \to y$ as $n \to \infty$, by Proposition 1.39 there exists $M > 0$ such that $|x_n| \leqslant M$ and $|y_n| \leqslant M$. Let $\varepsilon > 0$ be given. Then

$$(\exists\, N_1 \in \mathbb{N})\big(n \geqslant N_1 \Rightarrow |x_n - x\,| < \frac{\varepsilon}{2M}\big),$$

and

$$(\exists\, N_2 \in \mathbb{N})\big(n \geqslant N_2 \Rightarrow |y_n - y\,| < \frac{\varepsilon}{2M}\big).$$

Define $N = \max\{N_1, N_2\}$. Then $N \in \mathbb{N}$, and if $n \geqslant N$,

$$\begin{aligned}|x_n \cdot y_n - x \cdot y| &= |x_n \cdot y_n - x_n \cdot y + x_n \cdot y - x \cdot y| \leqslant |x_n \cdot (y_n - y)| + |y \cdot (x_n - x)| \\ &\leqslant M \cdot |y_n - y| + M \cdot |x_n - x| < M \cdot \frac{\varepsilon}{2M} + M \cdot \frac{\varepsilon}{2M} = \varepsilon.\end{aligned}$$

3. It suffices to show that $\lim\limits_{n\to\infty} \dfrac{1}{y_n} = \dfrac{1}{y}$ if $y_n, y \neq 0$ (because of 2). Since $\lim\limits_{n\to\infty} y_n = y$, there exists $N_1 \in \mathbb{N}$ such that $|y_n - y| < \dfrac{|y|}{2}$ whenever $n \geqslant N_1$. Therefore, $|y| - |y_n| < \dfrac{|y|}{2}$ for all $n \geqslant N_1$ which further implies that $|y_n| > \dfrac{|y|}{2}$ for all $n \geqslant N_1$.

Let $\varepsilon > 0$ be given. Since $\lim\limits_{n\to\infty} y_n = y$, there exists $N_2 \in \mathbb{N}$ such that $|y_n - y| < \dfrac{|y|^2}{2}\varepsilon$ whenever $n \geqslant N_2$. Define $N = \max\{N_1, N_2\}$. Then $N \in \mathbb{N}$ and if $n \geqslant N$,

$$\Big|\frac{1}{y_n} - \frac{1}{y}\Big| = \frac{|y_n - y|}{|y_n||y|} < \frac{|y|^2}{2}\varepsilon \cdot \frac{1}{|y|}\frac{2}{|y|} = \varepsilon. \qquad \square$$

Theorem 1.41. *An ordered field* $(\mathbb{F}, +, \cdot, \leqslant)$ *has Archimedean Property if and only if the sequence* $\Big\{\dfrac{1}{n}\Big\}_{n=1}^{\infty}$ *converges to* 0.

Proof. "$\Rightarrow$" Let $\varepsilon > 0$ be given. Define $x = \dfrac{1}{\varepsilon}$. Then $x > 0$ by Proposition 1.13. Moreover, Archimedean Property of $\mathbb{F}$ implies that there exists N such that $x < N$. Then $N > 0$ and if $n \geqslant N$,

$$\Big|\frac{1}{n} - 0\Big| = \frac{1}{n} \leqslant \frac{1}{N} < \frac{1}{x} = \varepsilon.$$

"$\Leftarrow$" Let $x \in \mathbb{F}$ be given. If $x \leqslant 0$, we choose $n = 1$ so that $x < 1$. If $x > 0$, let $\varepsilon = \dfrac{1}{x}$. Then $\varepsilon > 0$ by Proposition 1.13. Since $\dfrac{1}{n} \to 0$ as $n \to \infty$, there exists $N > 0$ such that

$$\frac{1}{n} = \left|\frac{1}{n} - 0\right| < \varepsilon = \frac{1}{x} \qquad \text{whenever} \quad n \geqslant N\,.$$

In particular, $\dfrac{1}{N} < \dfrac{1}{x}$ which implies that $x < N$. □

Remark 1.42. There are ordered fields that do not have Archimedean Property, and these fields are called non-Archimedean ordered fields (while an ordered field satisfying Archimedean Property is called Archimedean ordered fields - **AP** 有序體). In a non-Archimedean ordered field, the sequence $\left\{\dfrac{1}{n}\right\}_{n=1}^{\infty}$ does not converge to 0.

1.3 Monotone Sequence Property

Definition 1.43. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, and $\{x_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{F}$.

1. $\{x_n\}_{n=1}^{\infty}$ is said to be ***increasing/non-decreasing*** if $x_n \leqslant x_{n+1}$ for all $n \in \mathbb{N}$.
2. $\{x_n\}_{n=1}^{\infty}$ is said to be ***decreasing/non-increasing*** if $x_n \geqslant x_{n+1}$ for all $n \in \mathbb{N}$.
3. $\{x_n\}_{n=1}^{\infty}$ is said to be ***strictly increasing*** if $x_n < x_{n+1}$ for all $n \in \mathbb{N}$.
4. $\{x_n\}_{n=1}^{\infty}$ is said to be ***strictly decreasing*** if $x_n > x_{n+1}$ for all $n \in \mathbb{N}$.

A sequence is called (strictly) ***monotone*** if it is either (strictly) increasing or (strictly) decreasing.

Definition 1.44. An ordered field $\mathbb{F}$ is said to satisfy the ***monotone sequence property*** (**MSP**) if every bounded monotone sequence converges to a limit in $\mathbb{F}$.

Remark 1.45. An equivalent definition of the monotone sequence property is that every monotone *increasing* sequence *bounded from above* converges; that is, if each sequence $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{F}$ satisfying

(i) $x_n \leqslant x_{n+1}$ for all $n \in \mathbb{N}$,

(ii) there exists $M \in \mathbb{F}$ such that $x_n \leqslant M$ for all $n \in \mathbb{N}$,

is convergent, then we say $\mathbb{F}$ satisfies the monotone sequence property.

Example 1.46. Consider the sequence $\{y_n\}_{n=1}^{\infty}$ in $\mathbb{Q}$ defined by

$$y_1 = \frac{1}{2}, \qquad y_{n+1} = \frac{1}{2 + \dfrac{1}{2 + y_n}}.$$

1. $\{y_n\}_{n=1}^{\infty}$ is bounded from below by zero.

2. $\{y_n\}_{n=1}^{\infty}$ is a decreasing sequence in $\mathbb{Q}$ (which can be proved by induction).

If $\lim\limits_{n\to\infty} y_n = y$, then Theorem 1.40 implies that $y = \dfrac{1}{2 + \dfrac{1}{2+y}}$ from which we conclude that $y = -1 + \sqrt{2}$. Since $y \notin \mathbb{Q}$, $\{y_n\}_{n=1}^{\infty}$ does not converge (to a limit) in $\mathbb{Q}$. In other words, $\mathbb{Q}$ does not satisfy the monotone sequence property.

Proposition 1.47. *An ordered field satisfying the monotone sequence property satisfies Archimedean Property; that is, if $\mathbb{F}$ is an ordered field satisfying the monotone sequence property, then for all $x \in \mathbb{F}$, there exists $n \in \mathbb{N}$ such that $x < n$.*

Proof. Assume the contrary that there exists $x \in \mathbb{F}$ such that $n \leqslant x$ for all $n \in \mathbb{N}$. Let $x_n = n$. Then $\{x_n\}_{n=1}^{\infty}$ is increasing and bounded from above. By the monotone sequence property of $\mathbb{F}$, there exists $\widehat{x} \in \mathbb{F}$ such that $x_n \to \widehat{x}$ as $n \to \infty$; thus there exists $N > 0$ such that

$$|x_n - \widehat{x}| < \frac{1}{4} \qquad \text{whenever} \quad n \geqslant N.$$

In particular, $|N - \widehat{x}| < \frac{1}{4}$, $|N + 1 - \widehat{x}| < \frac{1}{4}$; thus

$$1 = |N + 1 - N| \leqslant |N + 1 - \widehat{x}| + |\widehat{x} - N| < \frac{1}{4} + \frac{1}{4} = \frac{1}{2},$$

a contradiction. □

Example 1.48. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field satisfying the monotone sequence property, and $y \in \mathbb{F}$ be a given positive number (that is, $y > 0$). Define $x_n = \dfrac{N_n}{2^n}$, where N_n is the largest integer such that $x_n^2 \leqslant y$; that is, $\left(\dfrac{N_n}{2^n}\right)^2 \leqslant y$ but $\left(\dfrac{N_n + 1}{2^n}\right)^2 > y$ (for example, if $y = 2$, then $x_1 = \dfrac{2}{2^1}$, $x_2 = \dfrac{5}{2^2}$, $x_3 = \dfrac{11}{2^3}$, $\cdots$). Then

1. x_n is bounded from above: since $x_n^2 \leqslant y \leqslant 2y + y^2 + 1 = (y+1)^2$, by the non-negativity of x_n and y and Remark 1.15 we must have $0 \leqslant x_n \leqslant y + 1$.

2. x_n is increasing: by the definition of N_n,

$$N_n^2 \leqslant 2^{2n} \cdot y \Rightarrow 4 \cdot N_n^2 \leqslant 2^{2n+2} \cdot y = 2^{2(n+1)} \cdot y \Rightarrow \Big(\frac{2N_n}{2^{n+1}}\Big)^2 \leqslant y \Rightarrow 2N_n \leqslant N_{n+1}\,.$$

Therefore, $x_n = \dfrac{N_n}{2^n} = \dfrac{2N_n}{2^{n+1}} \leqslant \dfrac{N_{n+1}}{2^{n+1}} = x_{n+1}$. Since $\mathbb{F}$ satisfies the monotone sequence property, there exists $x \in \mathbb{F}$ such that $x_n \to x$ as $n \to \infty$. By Theorem 1.40, $x_n^2 \to x^2$, and by Proposition 1.35, $x^2 \leqslant y$.

Now we show $x^2 = y$. To this end observe that

$$\Big(x_n + \frac{1}{2^n}\Big)^2 = \Big(\frac{N_n}{2^n} + \frac{1}{2^n}\Big) = \Big(\frac{N_n + 1}{2^n}\Big)^2 > y;$$

thus $x_n^2 \leqslant y \leqslant \big(x_n + \dfrac{1}{2^n}\big)^2$. By Archimedean propery of $\mathbb{F}$ (Proposition 1.47), $\lim\limits_{n\to\infty} \dfrac{1}{2^n} = 0$; thus Theorem 1.40 implies that $x^2 = \lim\limits_{n\to\infty} x_n^2 = \lim\limits_{n\to\infty} \big(x_n + \dfrac{1}{2^n}\big)^2 = y$. Note that Proposition 1.14 implies that such an x is unique if $x > 0$.

In general, one can define the n-th root of non-negative number y in an ordered field satisfying the monotone sequence property. The construction of the n-th root of $y \in \mathbb{F}$ is left as an exercise.

Definition 1.49. For $n \in \mathbb{N}$, the ***n-th root*** of a non-negative number y in an ordered field satisfying the monotone sequence property is the unique non-negative number x satisfying $x^n = y$. One writes $y^{1/n}$ or $\sqrt[n]{y}$ to denote n-th root of y.

1.4 Least Upper Bound Property

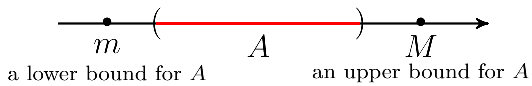
Definition 1.50. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, and $\varnothing \neq A \subseteq \mathbb{F}$. A number $M \in \mathbb{F}$ is called an ***upper bound*** (上界) for A if $x \leqslant M$ for all $x \in A$, and a number $m \in \mathbb{F}$ is called a ***lower bound*** (下界) for A if $x \geqslant m$ for all $x \in A$. If there is an upper bound for A, then A is said to be ***bounded from above***, while if there is a lower bound for A, then A is said to be ***bounded from below***. A number $b \in \mathbb{F}$ is called a ***least upper bound*** (最小上界) of A if

1. b is an upper bound for A, and

2. if M is an upper bound for A, then $M \geqslant b$.

A number a is called a ***greatest lower bound*** (最大下界) of A if

1. a is a lower bound for A, and
2. if m is a lower bound for A, then $m \leqslant a$.



If A is not bounded from above, the least upper bound of A is set to be ∞, while if A is not bounded from below, the greatest lower bound of A is set to be $-\infty$. The least upper bound of A is also called the ***supremum*** of A and is usually denoted by lub$A$ or $\sup A$, and "the" greatest lower bound of A is also called the ***infimum*** of A, and is usually denoted by glb$A$ or $\inf A$. If $A = \varnothing$, then $\sup A = -\infty$, $\inf A = \infty$.

We emphasize that "$\sup A = \infty$" is purely a notation denoting that A is not bounded from above; however, $\infty \notin \mathbb{F}$ and $\sup A$ does not exist.

Remark 1.51. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field.

1. If $b_1, b_2 \in \mathbb{F}$ are least upper bounds for a set $A \subseteq \mathbb{F}$, then $b_1 = b_2$ (since $b_1 \leqslant b_2$ and $b_2 \leqslant b_1$ for b_1, b_2 are also upper bounds for A). Therefore, $\sup A$ is a well-defined concept. Similarly, $\inf A$ is a well-defined concept.
2. Since the sentence "$x \in \varnothing \Rightarrow x \leqslant M$" is true for all $M \in \mathbb{F}$, we conclude that $\sup \varnothing = -\infty$. Similarly, $\inf \varnothing = \infty$.

Example 1.52. In the ordered field $\mathbb{R}$ (pretended that you know what $\mathbb{R}$ is),

1. $\sup(0,3) = 3$ and $\inf(0,3) = 0$.
2. $\sup \mathbb{N}$ does not exist, but $\inf \mathbb{N} = 1$.
3. Let $A = \{2^{-k} \,|\, k \in \mathbb{N}\}$. Then $\inf A = 0$ and $\sup A = \dfrac{1}{2}$.
4. Let $B = \{x \in \mathbb{Q} \,|\, x^2 < 2\}$. Then $\inf B = -\sqrt{2}$ and $\sup B = \sqrt{2}$.

How about considering the supremum and infimum for the sets above in the ordered field $\mathbb{Q}$?

Proposition 1.53. *Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, and A be a non-empty subset of $\mathbb{F}$.*

1. $b = \sup A \in \mathbb{F}$ *if and only if*

   (i) *b is an upper bound for A.* (ii) $(\forall\, \varepsilon > 0)(\exists\, x \in A)(x > b - \varepsilon)$.

2. $a = \inf A \in \mathbb{F}$ *if and only if*

   (i) *a is a lower bound for A.* (ii) $(\forall\, \varepsilon > 0)(\exists\, x \in A)(x < a + \varepsilon)$.

Proof. It suffices to prove 1.

"$\Rightarrow$" (i) is part of the definition of being a least upper bound.

(ii) If M is an upper bound for A, then we must have $M \geqslant b$; thus $b - \varepsilon$ is not an upper bound for A. Therefore, there exists $x \in A$ such that $x > b - \varepsilon$.

"$\Leftarrow$" We show that if M is an upper bound of A, then $M \geqslant b$. Assume the contrary that there exists an upper bound M for A satisfying $M < b$. Let $\varepsilon = b - M$. Then $\varepsilon > 0$ and there is no $x \in A$ satisfying $x > b - \varepsilon$, a contradiction. ▫

Corollary 1.54. *Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, and A be a non-empty subset of $\mathbb{F}$.*

1. $b = \sup A \in \mathbb{F}$ *if and only if*

   (i) $(\forall\, \varepsilon > 0)(\forall\, x \in A)(x < b + \varepsilon)$. (ii) $(\forall\, \varepsilon > 0)(\exists\, x \in A)(x > b - \varepsilon)$.

2. $a = \inf A \in \mathbb{F}$ *if and only if*

   (i) $(\forall\, \varepsilon > 0)(\forall\, x \in A)(x > a - \varepsilon)$. (ii) $(\forall\, \varepsilon > 0)(\exists\, x \in A)(x < a + \varepsilon)$.

Proof. By Proposition 1.53, it suffices to show that

$$\text{condition 1(i)} \quad \Leftrightarrow \quad b \text{ is an upper bound for } A.$$

Since the direction "$\Leftarrow$" is trivial, we only need to prove the direction "$\Rightarrow$". Suppose the contrary that b is not an upper bound for A. Then there exists $x \in A$ such that $b < x$. Let $\varepsilon = x - s$. Then $\varepsilon > 0$ and we do not have 1(i) since $x \in A$ but $x \nless s + \varepsilon$. ▫

Proposition 1.55. *Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an order field, and $\varnothing \neq A \subseteq B \subseteq \mathbb{F}$. Then* $\inf B \leqslant \inf A \leqslant \sup A \leqslant \sup B$ *whenever those numbers exist in $\mathbb{F}$ or are $\pm\infty$.*

Proof. We proceed as follows.

1. $\sup A \leqslant \sup B$: Let $b = \sup B$, then for all $x \in B$, $x \leqslant b$. Since $A \subseteq B$, then for all $x \in A$, $x \leqslant b$; hence b is also an upper bound for A. Since $\sup A$ is the least upper bound of A and b is an upper bound for A, then $\sup A \leqslant b = \sup B$.

2. It is similar to prove $\inf B \leqslant \inf A$.

3. It is trivially true that $\inf A \leqslant \sup A$. □

Definition 1.56 (Least Upper Bound Property)**.** Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field. $\mathbb{F}$ is said to satisfy the ***least upper bound property*** (**LUBP**) if every non-empty subset of $\mathbb{F}$ that has an upper bound in $\mathbb{F}$ has a supremum that is an element of $\mathbb{F}$（非空有上界的集合必有最小上界）. The ***greatest lower bound property*** (**GLBP**) for ordered fields is defined similarly.

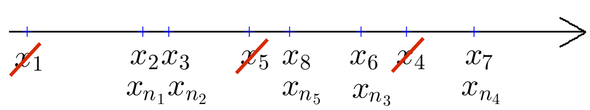
Proposition 1.57. *Every ordered field satisfying the least upper bound property satisfies Archimedean Property; that is, if $\mathbb{F}$ is an ordered field satisfying the least upper bound property, then for all $x \in \mathbb{F}$, there exists $n \in \mathbb{N}$ such that $x < n$.*

Proof. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field with the least upper bound property, and $x \in \mathbb{F}$ be given. If $x < 1$, then the choice $n = 1$ validates $n > x$. Suppose $x \geqslant 1$. Define $A = \{n \in \mathbb{N} \,\big|\, n \leqslant x\}$. Then $1 \in A$ and x is an upper bound for A. By the least upper bound property of $\mathbb{F}$, $s \equiv \sup A \in \mathbb{F}$ exists. Since s is the least upper bound of A, $s - 1$ is not an upper bound for A; thus there exists $m \in A$ such that $m > s - 1$ or $s < m + 1$. Then $m + 1 \notin A$ which implies that $m + 1 \nleqslant x$. The choice $n = m + 1$ then satisfies $n > x$. □

1.5 Bolzano-Weierstrass Property

Definition 1.58. A sequence $\{y_j\}_{j=1}^{\infty}$ is called a ***subsequence*** (子數列) of a sequence $\{x_n\}_{n=1}^{\infty}$ if there exists a strictly increasing function $\phi : \mathbb{N} \to \mathbb{N}$ such that $y_j = x_{\phi(j)}$. In this case, we often write $\phi(j) = n_j$ and $y_j = x_{n_j}$.

In other words, a subsequence is a sequence that can be derived from another sequence by deleting some elements without changing the order of remaining elements. Let $f : \mathbb{N} \to \mathbb{F}$ be a sequence and $x_n = f(n)$. A subsequence $\{x_{n_j}\}_{j=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ is the image of an infinite subset $\{n_1, n_2, \cdots\}$ of $\mathbb{N}$ under the map f (or simply the sequence $f \circ \phi : \mathbb{N} \to \mathbb{F}$).



Example 1.59. Consider the sequence $\{x_n\}_{n=1}^{\infty}$ defined recursively by

$$x_1 = \frac{1}{2}, \qquad x_2 = \frac{1}{2+\frac{1}{2}}, \qquad \cdots, \qquad x_{n+1} = \frac{1}{2+x_n}.$$

Then the sequence $\{y_n\}_{n=1}^{\infty}$ given in Example 1.46 is a subsequence of $\{x_n\}_{n=1}^{\infty}$. In fact,

$$y_n = x_{2n-1} \quad \forall\, n \in \mathbb{N} \qquad (\text{with the choice of } \phi(n) = 2n-1).$$

The following proposition concerns equivalent conditions for the convergence of sequences.

Proposition 1.60. *Let* $(\mathbb{F}, +, \cdot, \leqslant)$ *be an ordered field,* $\{x_n\}_{n=1}^{\infty}$ *be a sequence in* $\mathbb{F}$*, and* $x \in \mathbb{F}$*. Then*

1. $x_n \to x$ *as* $n \to \infty$ *if and only if every proper subsequence of* $\{x_n\}_{n=1}^{\infty}$ *converges to* x*.*
2. $x_n \to x$ *as* $n \to \infty$ *if and only if every proper subsequence of* $\{x_n\}_{n=1}^{\infty}$ *has a further subsequence that converges to* x*.*

Proof. 1. "$\Rightarrow$" Let $\{x_{n_j}\}_{j=1}^{\infty}$ be a subsequence of a convergent sequence $\{x_n\}_{n=1}^{\infty}$ with limit x, and $\varepsilon > 0$ be given. Then there exists $N > 0$ such that $|x_n - x| < \varepsilon$ whenever $n \geqslant N$. Since $n_j \geqslant j$ for all $j \in \mathbb{N}$, we find that $|x_{n_j} - x| < \varepsilon$ whenever $j \geqslant N$.

"$\Leftarrow$" Let $\varepsilon > 0$ be given. Since every proper subsequence of $\{x_n\}_{n=1}^{\infty}$ converges to x, the subsequence $\{x_{n+1}\}_{n=1}^{\infty}$ converges to x. Therefore, there exists $N_1 > 0$ such that

$$|x_{n+1} - x| < \varepsilon \qquad \text{whenever} \quad n \geqslant N_1.$$

Let $N = N_1 + 1$. Then $|x_n - x| < \varepsilon$ whenever $n \geqslant N$.

2. The direction "$\Rightarrow$" follows from 1. For the direction "$\Leftarrow$", assume the contrary that $x_n \nrightarrow x$ as $n \to \infty$. Then

$$(\exists\, \varepsilon > 0)\big(\#\big\{n \in \mathbb{N} \,\big|\, |x_n - x| \geqslant \varepsilon\big\} = \infty\big).$$

Let $\big\{n \in \mathbb{N} \,\big|\, |x_n - x| \geqslant \varepsilon\big\} = \big\{n_j \in \mathbb{N} \,\big|\, n_j < n_{j+1} \text{ for all } j \in \mathbb{N}\big\}$. The subsequence $\{x_{n_j}\}_{j=1}^{\infty}$ clearly does not have any subsequence $\big\{x_{n_{j_k}}\big\}_{k=1}^{\infty}$ which converge to x, a contradiction. $\square$

Remark 1.61. 1 of Proposition 1.60 indeed can be rephrased as "$\{x_n\}_{n=1}^\infty$ converges if and only if every proper subsequence of $\{x_n\}_{n=1}^\infty$ converges". This fact is left as an exercise.

Recall that sequence $\{x_n\}_{n=1}^\infty$ converges to x if and only if

$$\forall\, \varepsilon > 0,\ \#\big\{n \in \mathbb{N} \,\big|\, x_n \notin (x-\varepsilon, x+\varepsilon)\big\} < \infty\,.$$

The statement above implies that

$$\forall\, \varepsilon > 0,\ \#\big\{n \in \mathbb{N} \,\big|\, x_n \in (x-\varepsilon, x+\varepsilon)\big\} = \infty\,; \tag{1.5.1}$$

however, if x satisfies (1.5.1), x might not be the limit of the sequence. Nevertheless, a candidate for the limit of a sequence must satisfy (1.5.1), and we call such a point a cluster point of $\{x_n\}_{n=1}^\infty$. To be more precise, we have the following

Definition 1.62. A point x is called a ***cluster point*** of a sequence $\{x_n\}_{n=1}^\infty$ if (1.5.1) holds.

We note that (1.5.1) is equivalent to that

$$\forall\, \varepsilon > 0,\ \#\big\{n \in \mathbb{N} \,\big|\, |x_n - x| < \varepsilon\big\} = \infty\,.$$

Example 1.63. Let $x_n = (-1)^n$. Then 1 and -1 are the only two cluster points of $\{x_n\}_{n=1}^\infty$.

Example 1.64. Let $x_n = (-1)^n + \dfrac{1}{n}$.
Claim: 1 and -1 are cluster points of $\{x_n\}_{n=1}^\infty$.
Let $\varepsilon > 0$ be given. We observe that

$$\big\{n \in \mathbb{N} \,\big|\, x_n \in (1-\varepsilon, 1+\varepsilon)\big\} \supseteq \Big\{n \in \mathbb{N} \,\Big|\, n \text{ is even}, \frac{1}{n} < \varepsilon\Big\}\,;$$

thus $\#\big\{n \in \mathbb{N} \,\big|\, x_n \in (1-\varepsilon, 1+\varepsilon)\big\} = \infty$. Similarly, -1 is a cluster point. Moreover, if $a \neq \pm 1$, a is not a cluster point of $\{x_n\}_{n=1}^\infty$.

Example 1.65. Let $S = \mathbb{Q} \cap [0,1]$. Then S is countable since it is a subset of a countable set $\mathbb{Q}$. Therefore, there exists $f : \mathbb{N} \xrightarrow[\text{onto}]{1-1} S$ or equivalently $S = \{q_1, q_2, \cdots, q_n, \cdots\}$. The collection of all cluster points of $\{q_n\}_{n=1}^\infty$ is $[0,1]$ since every open interval (with mid-point in $[0,1]$) contains infinitely many rational numbers in S.

Definition 1.66 (Bolzano-Weierstrass Property)**.** Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field. $\mathbb{F}$ is said to satisfy the ***Bolzano-Weierstrass property*** (**BWP**) if every bounded sequence in $\mathbb{F}$ has a convergent subsequence; that is, every bounded sequence in $\mathbb{F}$ has a subsequence that converges to a limit in $\mathbb{F}$.

Remark 1.67. $\mathbb{Q}$ does not satisfy the Bolzano-Weierstrass property. Example 1.46 provides an counterexample of a bounded divergent sequence in $\mathbb{Q}$.

Proposition 1.68. *Let $(\mathbb{F},+,\cdot,\leqslant)$ be an ordered field satisfying the Bolzano-Weierstrass Property, $\{x_n\}_{n=1}^\infty$ be a sequence in $\mathbb{F}$, and $x \in \mathbb{F}$. Then $x_n \to x$ as $n \to \infty$ if and only if $\{x_n\}_{n=1}^\infty$ is bounded and x is the only cluster point of $\{x_n\}_{n=1}^\infty$.*

Proof. "$\Rightarrow$" The boundedness is concluded by Proposition 1.39. For uniqueness, suppose the contrary that there is another cluster point $y \neq x$. Let $\varepsilon = \dfrac{|x-y|}{2}$. Then

$$\#\big\{n \in \mathbb{N} \,\big|\, |x_n - y| < \varepsilon\big\} = \infty\,.$$

Since $\big\{n \in \mathbb{N} \,\big|\, |x_n - x| > \varepsilon\big\} \supseteq \big\{n \in \mathbb{N} \,\big|\, |x_n - y| < \varepsilon\big\}$ (this inclusion is left as an exercise), we find that

$$\#\big\{n \in \mathbb{N} \,\big|\, |x_n - x| > \varepsilon\big\} = \infty\,,$$

a contradiction to that $x_n \to x$ as $n \to \infty$.

"$\Leftarrow$" Suppose that $\{x_n\}_{n=1}$ is a bounded sequence in $\mathbb{F}$ and has x as the only cluster point but $\{x_n\}_{n=1}^\infty$ does not converge to x. Then

$$\exists\, \varepsilon > 0 \ni \#\big\{n \in \mathbb{N} \,\big|\, x_n \notin (x-\varepsilon, x+\varepsilon)\big\} = \infty\,.$$

Write $\big\{n \in \mathbb{N} \,\big|\, x_n \notin (x-\varepsilon, x+\varepsilon)\big\} = \{n_1, n_2, \cdots, n_k, \cdots\}$. Then we find a subsequence $\big\{x_{n_k}\big\}_{k=1}^\infty$ lying outside $(x-\varepsilon, x+\varepsilon)$. Since $\big\{x_{n_k}\big\}_{k=1}^\infty$ is bounded, the Bolzano-Weierstrass Property implies that there exists a convergent subsequence $\big\{x_{n_{k_j}}\big\}_{j=1}^\infty$ with limit y. Since $x_{n_{k_j}} \notin (x-\varepsilon, x+\varepsilon)$, $y \notin (x-\varepsilon, x+\varepsilon)$ by Proposition 1.35; thus $y \neq x$. On the other hand, the limit $\lim\limits_{j\to\infty} x_{n_{k_j}} = y$ implies that for every $\varepsilon > 0$,

$$\big\{j \in \mathbb{N} \,\big|\, \big|x_{n_{k_j}} - y\big| < \varepsilon\big\} \supseteq \big\{j \in \mathbb{N} \,\big|\, j \geqslant J\big\}$$

for some $J > 0$; thus $\#\big\{j \in \mathbb{N} \,\big|\, \big|x_{n_{k_j}} - y\big| < \varepsilon\big\} = \infty$ which shows that y is also a cluster point of $\{x_n\}_{n=1}^\infty$, a contradiction to the assumption that x is the only cluster point of $\{x_n\}_{n=1}^\infty$. $\square$

Example 1.69. Consider the sequence $\{x_n\}_{n=1}^\infty$ defined by

$$x_n = \begin{cases} n & \text{if } n \text{ is odd}, \\ 1 & \text{if } n \text{ is even}. \end{cases}$$

Then 1 is the only cluster point of $\{x_n\}_{n=1}^\infty$, but $\{x_n\}_{n=1}^\infty$ does not converge to 1 (since x_n is not bounded).

1.6 Cauchy Sequences

If a sequence $\{x_n\}_{n=1}^\infty$ in an ordered field $\mathbb{F}$ converges, then

$$\exists!\, x \in \mathbb{F} \ni \forall\, \varepsilon > 0, \#\big\{n \in \mathbb{N} \,\big|\, x_n \notin (x-\varepsilon, x+\varepsilon)\big\} < \infty\,.$$

We note that the statement above implies that if $\{x_n\}_{n=1}^\infty$ converges, then

$$(\forall\, \varepsilon > 0)(\exists \text{ an interval } I \text{ of length } 2\varepsilon)\big(\#\big\{n \in \mathbb{N} \,\big|\, x_n \notin I\big\} < \infty\big)\,. \tag{$\star$}$$

The statement above motivates the following

Definition 1.70. A sequence $\{x_n\}_{n=1}^\infty$ in an ordered field is said to be ***Cauchy*** if

$$(\forall\, \varepsilon > 0)(\exists\, N > 0)\big(n, m \geqslant N \Rightarrow |x_n - x_m| < \varepsilon\big)\,.$$

Remark 1.71. ($\star$) 這個敘述的中心思想是：給定任一正數 ε, 我們都能找到一個長度是 2ε 的區間使得落在此區間外的 x_n 只有有限個。因為當對每個長度我們都能找到這樣的區間時，才有機會找到 $\{x_n\}_{n=1}^\infty$ 的極限（而這個極限一定落在所有這樣的區間之內）。

Example 1.72. In $\mathbb{Q}$, $x_1 = 3, x_2 = 3.1, x_3 = 3.14, x_4 = 3.141, \cdots$. Then $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence, but is not convergent. Therefore, a Cauchy sequence in an ordered field may not converge.

Example 1.73. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an Archimedean ordered field, and $\{x_n\}_{n=1}^\infty \subseteq \mathbb{F}$ be a sequence satisfying $|x_n - x_{n+1}| < \dfrac{1}{2^{n+1}}$ for all $n \in \mathbb{N}$.
Claim: $\{x_n\}_{n=1}^\infty$ is Cauchy. Given $\varepsilon > 0$, choose $N > 0$ such that $\dfrac{1}{2^N} < \varepsilon$ (such an N exists because of Theorem 1.41 and the fact that $2^N > N$ for all $N \in \mathbb{N}$). Then if $N \leqslant n < m$,

$$\begin{aligned}
|x_n - x_m| &\leqslant |x_n - x_{n+1}| + |x_{n+1} - x_m| \\
&\leqslant \cdots \\
&\leqslant |x_n - x_{n+1}| + |x_{n+1} - x_{n+2}| + \cdots + |x_{m-1} - x_m| \\
&\leqslant \frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \cdots + \frac{1}{2^m} \\
&\leqslant \frac{1}{2^n} \leqslant \frac{1}{2^N} < \varepsilon;
\end{aligned}$$

thus $\{x_n\}_{n=1}^\infty$ is Cauchy in $\mathbb{F}$.

Proposition 1.74. *Every convergent sequence is Cauchy.*

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a convergent sequence with limit x, and $\varepsilon > 0$ be given. By the definition of the convergence of sequences, there exists $N > 0$ such that $|x_n - x| < \dfrac{\varepsilon}{2}$ whenever $n \geqslant N$. Then by triangle inequality, if $n, m \geqslant N$,

$$|x_n - x_m| \leqslant |x_n - x| + |x - x_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon;$$

thus $\{x_n\}_{n=1}^{\infty}$ is Cauchy. □

Lemma 1.75. *Every Cauchy sequence is bounded.*

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be Cauchy. There exists $N > 0$ such that $|x_n - x_m| < 1$ for all $n, m \geqslant N$. In particular, $|x_n - x_N| < 1$ if $n \geqslant N$ or equivalently,

$$x_N - 1 < x_n < x_N + 1 \qquad \forall\, n \geqslant N.$$

Let $M = \max\big\{|x_1|, |x_2|, \cdots, |x_{N-1}|, |x_N| + 1\big\}$. Then $|x_n| \leqslant M$ for all $n \in \mathbb{N}$. □

Lemma 1.76. *If a subsequence of a Cauchy sequence is convergent, then this Cauchy sequence also converges.*

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence with a convergent subsequence $\{x_{n_j}\}_{j=1}^{\infty}$ whose limit is x, and $\varepsilon > 0$ be given. Then there exist $K, N > 0$ such that

$$|x_{n_j} - x| < \frac{\varepsilon}{2} \text{ whenever } j \geqslant K, \quad \text{and} \quad |x_n - x_m| < \frac{\varepsilon}{2} \text{ whenever } n, m \geqslant N.$$

Choose $j \geqslant \max\{K, N\}$. Then $n_j \geqslant N$; thus if $n \geqslant N$,

$$|x_n - x| \leqslant |x_n - x_{n_j}| + |x_{n_j} - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$ □

Remark 1.77. Combining Proposition 1.60 and Lemma 1.76, we conclude that a Cauchy sequence converges if and only if a subsequence converges.

1.7 Completeness

In this section, we establish the equivalency between those properties introduced in the previous sections. These equivalent properties lead to an important concept, the completeness of ordered fields. There is exactly one ordered field satisfying all these properties, and this ordered field will be the real number system $\mathbb{R}$.

Theorem 1.78. *An ordered field satisfies the monotone sequence property if and only if it satisfies the least upper bound property*（有序體中 **MSP** 與 **LUBP** 為等價性質）.

Proof. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field.

"$\Leftarrow$" Suppose that $\mathbb{F}$ satisfies the least upper bound property, and let $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{F}$ be an increasing sequence bounded from above. By the least upper bound property of $\mathbb{F}$, the set $\{x_1, x_2, \cdots, x_n, \cdots\}$ has a least upper bound $x \in \mathbb{F}$. We next show that x is the limit of $\{x_n\}_{n=1}^{\infty}$.

Let $\varepsilon > 0$ be given. By Corollary 1.54, there exists x_N such that $x_N > x - \varepsilon$. Therefore, the fact that $\{x_n\}_{n=1}^{\infty}$ is increasing and bounded from above by x imply that
$$x - \varepsilon < x_n \leqslant x < x + \varepsilon \qquad \forall\, n \geqslant N\,.$$
This shows that $|x_n - x| < \varepsilon$ whenever $n \geqslant N$; thus $x_n \to x$ as $n \to \infty$.

"$\Rightarrow$ Suppose that $\mathbb{F}$ satisfies the monotone sequence property, and let A be a non-empty subset of $\mathbb{F}$ bounded from above. Let N_n is the largest integer satisfying that $\dfrac{N_n}{2^n}$ is not an upper bound for A but $\dfrac{N_n + 1}{2^n}$ is an upper bound for A, and define $x_n = \dfrac{N_n}{2^n}$. If M is an upper bound for A, $x_n \leqslant M$ for all $n \in \mathbb{N}$; thus $\{x_n\}_{n=1}^{\infty}$ is bounded from above. Moreover, by the fact that N_{n+1} is the largest integer satisfying $\dfrac{N_{n+1}}{2^{n+1}}$ is not an upper bound for A, we must have
$$x_n = \frac{N_n}{2^n} = \frac{2N_n}{2^{n+1}} \leqslant \frac{N_{n+1}}{2^{n+1}} = x_{n+1}\,;$$
thus $\{x_n\}_{n=1}^{\infty}$ is an increasing sequence. Therefore, the monotone sequence property implies that $\{x_n\}_{n=1}^{\infty}$ converges to a limit $x \in \mathbb{F}$. Next we show that x is the least upper bound of A.

Let $\varepsilon > 0$ be given. Then $x + \varepsilon$ must be an upper bound for A for otherwise if $\varepsilon > \dfrac{1}{2^k}$ for some $k \in \mathbb{N}$, then N_k is not the largest integer satisfying the required property. On the other hand, since $x_n \to x$ as $n \to \infty$, there exists $N > 0$ such that $|x_n - x| < \varepsilon$ whenever $n \geqslant N$. Therefore, $x_N > x - \varepsilon$ which shows that $x - \varepsilon$ cannot be an upper bound for A; thus there exists $y \in A$ such that $y > x - \varepsilon$. By Corollary 1.54, we conclude that $x = \sup A$. ▫

Theorem 1.79. *An ordered field satisfying the monotone sequence property satisfies the Bolzano-Weierstrass property*（具 **MSP** 的有序體亦有 **BWP**）.

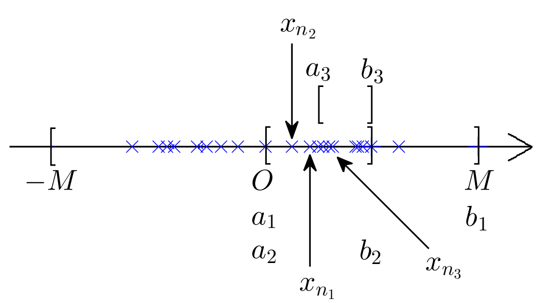
Proof. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field satisfying the monotone sequence property, and $\{x_n\}_{n=1}^{\infty}$ be a bounded sequence satisfying $|x_n| \leqslant M$ for all $n \in \mathbb{N}$. Divide $[-M, M]$ into two intervals $[-M, 0], [0, M]$, and denote one of the two intervals containing infinitely many x_n as $[a_1, b_1]$; that is, $\#\{n \in \mathbb{N} \,\big|\, x_n \in [a_1, b_1]\} = \infty$. Divide $[a_1, b_1]$ into two intervals $\big[a_1, \frac{a_1 + b_1}{2}\big], \big[\frac{a_1 + b_1}{2}, b_1\big]$, and denote one of the two intervals containing infinitely many x_n as $[a_2, b_2]$. We continue this process, and obtain a sequence of intervals $[a_k, b_k]$ such that $[a_{k+1}, b_{k+1}] \subseteq [a_k, b_k]$, $|b_k - a_k| = \dfrac{M}{2^{k-1}}$ and $\#\{n \in \mathbb{N} \,|\, x_n \in [a_k, b_k]\} = \infty$ for all $k \in \mathbb{N}$.

Since $[a_k, b_k] \supseteq [a_{k+1}, b_{k+1}]$ for all $k \in \mathbb{N}$, we find that $\{a_k\}_{k=1}^{\infty}$ is increasing and $\{b_k\}_{k=1}^{\infty}$ is decreasing. Moreover, $a_k \leqslant M$, $b_k \geqslant -M$. Therefore, the monotone sequence property implies that a_k converges to $a \in \mathbb{F}$ and b_k converges to $b \in \mathbb{F}$. On the other hand,

$$b - a = \lim_{k \to \infty} (b_k - a_k) = \lim_{k \to \infty} \frac{M}{2^{k-1}} = 0,$$

where we have used Proposition 1.47 along with Theorem 1.41 (and Theorem 1.40) to conclude the limit. Therefore, $a = b$.

Finally, we construct a convergent subsequence of $\{x_n\}_{n=1}^{\infty}$. Let x_{n_1} be an element belonging to $[a_1, b_1]$. Since $\#\{n \in \mathbb{N} \,|\, x_n \in [a_1, b_1]\} = \infty$, we can choose $n_2 > n_1$ such that $x_{n_2} \in [a_2, b_2]$, and for the same reason we can choose $n_3 > n_2$ such that $x_{n_3} \in [a_3, b_3]$. We continue this process and obtain a subsequence $x_{n_k} \in [a_k, b_k]$ with $n_k > n_{k-1}$.



Since $a_k \leqslant x_{n_k} \leqslant b_k$ for all $k \in \mathbb{N}$, by Sandwich Lemma $\lim\limits_{k \to \infty} x_{n_k} = a = b$. □

Theorem 1.80. *Every Cauchy sequence in an ordered field satisfying the Bolzano-Weierstrass property converges*（具 **BWP** 的有序體中的柯西數列必收斂）.

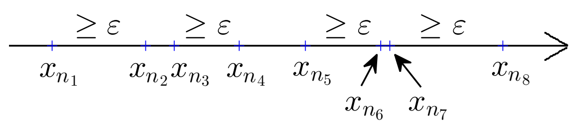
Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence in an ordered field satisfying the Bolzano-Weierstrass property. By Lemma 1.75, $\{x_n\}_{n=1}^{\infty}$ is bounded; thus the Bolzano-Weierstrass property provides a convergent subsequence $\{x_{n_j}\}_{j=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$. The convergence of $\{x_n\}_{n=1}^{\infty}$ is then guaranteed by Lemma 1.76. □

Theorem 1.81. *An Archimedean ordered field satisfying the property that every Cauchy sequence converges satisfies the monotone sequence property*（滿足柯西數列必收斂的 **AP** 有序體也有 **MSP**）.

Proof. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an Archimedean ordered field. Suppose the contrary that there is a bounded increasing sequence $\{x_n\}_{n=1}^{\infty}$ that does not converge to a limit in $\mathbb{F}$. By the assumption that every Cauchy sequence converges, $\{x_n\}_{n=1}^{\infty}$ cannot be Cauchy; thus

$$(\exists\, \varepsilon > 0)(\forall\, N > 0)(\exists\, n, m \geqslant N)(|x_n - x_m| \geqslant \varepsilon).$$

Let $N = 1$, there exists $n_2 > n_1 \geqslant 1$ such that $|x_{n_1} - x_{n_2}| \geqslant \varepsilon$. Let $N = n_2 + 1$, there exists $n_4 > n_3 \geqslant n_2 + 1$ such that $|x_{n_3} - x_{n_4}| \geqslant \varepsilon$. We continue this process and obtain a sequence $\{x_{n_j}\}_{j=1}^{\infty}$ satisfying $\big|x_{n_{2k-1}} - x_{n_{2k}}\big| \geqslant \varepsilon$ for all $k \in \mathbb{N}$.



Suppose that $\{x_n\}_{n=1}^{\infty}$ is bounded from above by M; that is, $x_n \leqslant M$ for all $n \in \mathbb{N}$. Then for each $k \in \mathbb{N}$,

$$\begin{aligned} M \geqslant x_{n_{2k}} &= x_{n_{2k}} - x_{n_{2k-1}} + x_{n_{2k-1}} \geqslant \varepsilon + x_{n_{2k-2}} = \varepsilon + x_{n_{2k-2}} - x_{n_{2k-3}} + x_{n_{2k-3}} \\ &\geqslant \varepsilon + \varepsilon + x_{n_{2k-4}} = 2\varepsilon + x_{n_{2k-4}} - x_{n_{2k-5}} + x_{n_{2k-5}} \geqslant \cdots \geqslant (k-1)\varepsilon + x_{n_1}; \end{aligned}$$

thus

$$k \leqslant 1 + \frac{M - x_{n_1}}{\varepsilon} \qquad \forall\, k \in \mathbb{N},$$

a contradiction to Archimedean Property. $\square$

Summary: In an Archimedean ordered field, the following four properties are equivalent:

1. the monotone sequence property (單調有界數列必收斂),
2. the least upper bound property (非空集合有上界必有最小上界),
3. the Bolzano-Weierstrass property (有界數列必有收斂子數列),
4. the property that every Cauchy sequence converges (柯西數列必收斂).

Such property is called the completeness（完備性）, and we have the following

Definition 1.82. An ordered field $\mathbb{F}$ is said to be ***complete*** (完備) (or have the completeness property, 具備完備性) if it satisfies the monotone sequence property.

Theorem 1.83. *There is a complete ordered field. Moreover, if* $(\mathbb{F}, +, \cdot, \leqslant)$ *and* $(\mathcal{F}, \oplus, \odot, \preccurlyeq)$ *are complete ordered fields, there exists a bijection* $\phi : \mathbb{F} \to \mathcal{F}$ *such that*

1. $\phi(x + y) = \phi(x) \oplus \phi(y)$ *and* $\phi(x \cdot y) = \phi(x) \odot \phi(y)$ *for all* $x, y \in \mathbb{F}$.

2. *The order* $\leqslant$ *and* $\preccurlyeq$ *is consistent under the map* ϕ*; that is, if* $x \leqslant y$*, then* $\phi(x) \preccurlyeq \phi(y)$.

In other words, two complete order fields are isomorphic (so that there is one and only one complete ordered field).

Axiom of the completeness of real number system $\mathbb{R}$: The real number system $\mathbb{R}$ is complete.

Theorem 1.84. *Every Cauchy sequence in* $\mathbb{R}$ *is convergent.*

Remark 1.85. Let $f : A \to \mathbb{R}$ be a real-valued function. Then the supremum of the image of A under f is denoted by $\sup\limits_A f$ or $\sup\limits_{x \in A} f(x)$. In other words,

$$\sup_A f = \sup_{x \in A} f(x) = \sup \big\{ f(x) \,\big|\, x \in A \big\}.$$

Similarly, the infimum of the image of A under f is denoted by $\inf\limits_A f$ or $\inf\limits_{x \in A} f(x)$.

1.8 Limit Inferior and Limit Superior

Definition 1.86. A sequence $\{x_n\}_{n=1}^{\infty}$ is said to ***diverge to infinity*** if for all $M > 0$, there exists $N > 0$ such that $x_n > M$ whenever $n \geqslant N$. It is said to ***diverge to negative infinity*** if $\{-x_n\}_{n=1}^{\infty}$ diverge to infinity. We use $\lim\limits_{n \to \infty} x_n = \infty$ or $-\infty$ to denote that $\{x_n\}_{n=1}^{\infty}$ diverges to infinity or negative infinity.

Remark 1.87. By Definition 1.26, the limit of a sequence $\{x_n\}_{n=1}^{\infty}$ does not exist if $\lim\limits_{n \to \infty} x_n = \infty$ or $-\infty$; however, we sometimes also call ∞ or $-\infty$ the limit of $\{x_n\}_{n=1}^{\infty}$.

Definition 1.88. The ***extended real number system***, denoted by $\mathbb{R}^*$, is the number system $\mathbb{R} \cup \{\infty, -\infty\}$, where ∞ and $-\infty$ are two symbols satisfying $-\infty < x < \infty$ for all $x \in \mathbb{R}$.

Remark 1.89. 1. $\mathbb{R}^*$ is not a field since ∞ and $-\infty$ do not have multiplicative inverse.

2. The definition of the least upper bound of a set can be simplified as follows: Let $S \subseteq \mathbb{R}^*$ be a set (not necessary non-empty set). A number $b \in \mathbb{R}^*$ is said to be the least upper bound of S if

   (a) b is an upper bound for S (that is, $s \leqslant b$ for all $s \in S$);

   (b) If $M \in \mathbb{R}^*$ is an upper bound for S, then $b \leqslant M$.

   No further discussion (such as $S = \varnothing$ or S is not bounded from above) has to be made. The greatest lower bound can be defined in a similar fashion.

3. Any sets in $\mathbb{R}^*$ has a least upper bound and a greatest lower bound in $\mathbb{R}^*$, even the empty set and unbounded set.

4. Proposition 1.53 for the case $\mathbb{F} = \mathbb{R}$ can be rephrased as follows: Let $S \subseteq \mathbb{R}^*$. Then $b = \sup S \in \mathbb{R}$ if and only if

   (a) b is an upper bound for S;

   (b) for all $\varepsilon > 0$, there exists $s \in S$ such that $s > b - \varepsilon$.

   Note that $b \in \mathbb{R}$ is crucial since there is no $s \in \mathbb{R}^*$ such that $s > \infty - \varepsilon = \infty$. The greatest lower bound counterpart can be made in a similar fashion.

5. In light of Definition 1.86, we can redefine cluster points of a real sequence as follows: A number $x \in \mathbb{R}^*$ is said to be a cluster point of a sequence $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ if there exists a subsequence $\{x_{n_j}\}_{j=1}^{\infty}$ such that $\lim\limits_{j\to\infty} x_{n_j} = x$. Note that now we can talk about if ∞ or $-\infty$ is a cluster points of a real sequence.

In the rest of the section, one is allowed to find the least upper bound and the greatest lower bound of a subset in $\mathbb{R}^*$.

Definition 1.90. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{R}$.

1. The ***limit superior*** of $\{x_n\}_{n=1}^{\infty}$, denoted by $\limsup\limits_{n\to\infty} x_n$ or $\overline{\lim\limits_{n\to\infty}}\, x_n$, is the infimum of the set $\Big\{ \sup \{x_n \,\big|\, n \geqslant k\} \,\Big|\, k \in \mathbb{N} \Big\}$.

2. The ***limit inferior*** of $\{x_n\}_{n=1}^{\infty}$, denoted by $\liminf\limits_{n\to\infty} x_n$ or $\underline{\lim\limits_{n\to\infty}}\, x_n$, is the supremum of the set $\Big\{ \inf \{x_n \,\big|\, n \geqslant k\} \,\Big|\, k \in \mathbb{N} \Big\}$.

Remark 1.91. Let $\sup\limits_{n\geqslant k} x_n$ denote the number $\sup\big\{x_n \,\big|\, n \geqslant k\big\}$ and $\inf\limits_{n\geqslant k} x_n$ denote the number $\inf\big\{x_n \,\big|\, n \geqslant k\big\}$. Then the limit superior and the limit inferior can be written as

$$\limsup_{n\to\infty} x_n = \inf_{k\geqslant 1}\sup_{n\geqslant k} x_n \quad\text{and}\quad \liminf_{n\to\infty} x_n = \sup_{k\geqslant 1}\inf_{n\geqslant k} x_n\,.$$

Remark 1.92. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{R}$, and $y_k = \sup\limits_{n\geqslant k} x_n$ and $z_k = \inf\limits_{n\geqslant k} x_n$. Then $\{y_k\}_{k=1}^{\infty}$ is a decreasing sequence, and $\{z_k\}_{k=1}^{\infty}$ is an increasing sequence. Therefore, the limit of $\{y_k\}_{k=1}^{\infty}$ and the limit of $\{z_k\}_{k=1}^{\infty}$ both "exist" in the sense of Definition 1.26 and Remark 1.87. In fact, the limit of $\{y_k\}_{k=1}^{\infty}$ is the infimum of $\{y_k\}_{k=1}^{\infty}$, and the limit of $\{z_k\}_{k=1}^{\infty}$ is the supremum of $\{z_k\}_{k=1}^{\infty}$. In other words,

$$\lim_{k\to\infty}\sup_{n\geqslant k} x_n = \inf_{k\geqslant 1}\sup_{n\geqslant k} x_n \quad\text{and}\quad \lim_{k\to\infty}\inf_{n\geqslant k} x_n = \sup_{k\geqslant 1}\inf_{n\geqslant k} x_n\,;$$

thus

$$\limsup_{n\to\infty} x_n = \lim_{k\to\infty}\sup_{n\geqslant k} x_n \quad\text{and}\quad \liminf_{n\to\infty} x_n = \lim_{k\to\infty}\inf_{n\geqslant k} x_n\,.$$

Example 1.93. Let $\{x_n\}_{n=1}^{\infty}$ be the sequence given by $x_n = (-1)^n$. Then

$$y_k = \sup_{n\geqslant k} x_n = 1 \quad\text{and}\quad z_k = \inf_{n\geqslant k} x_n = -1\,.$$

Therefore, $\limsup\limits_{n\to\infty} x_n = \lim_{k\to\infty} y_k = 1$ and $\liminf\limits_{n\to\infty} x_n = \lim\limits_{k\to\infty} z_k = -1$.

Example 1.94. Let $\{x_n\}_{n=1}^{\infty}$ be a real sequence given by $x_n = \dfrac{1}{n}$. Then

$$y_k = \sup_{n\geqslant k} x_n = \frac{1}{k} \quad\text{and}\quad z_k = \inf_{n\geqslant k} x_n = 0\,.$$

Therefore, $\limsup\limits_{n\to\infty} x_n = \lim\limits_{k\to\infty} y_k = 0$ and $\liminf\limits_{n\to\infty} x_n = \lim\limits_{k\to\infty} z_k = 0$.

Example 1.95. Let $x_n = \begin{cases} 0 & \text{if } n \text{ is even} \\ n & \text{if } n \text{ is odd} \end{cases}$; that is, $\{x_n\}_{n=1}^{\infty} = \{1, 0, 3, 0, 5, \cdots\}$. Then

$$y_k = \sup_{n\geqslant k} x_n \quad\text{and}\quad z_k = \inf_{n\geqslant k} x_n = 0\,.$$

Therefore, $\limsup\limits_{n\to\infty} x_n = \lim\limits_{k\to\infty} y_k = \infty$ and $\liminf\limits_{n\to\infty} x_n = \lim\limits_{k\to\infty} z_k = 0$.

Example 1.96. Let $\{x_n\}_{n=1}^{\infty}$ be a real sequence defined by $x_n = (-1)^n + \dfrac{1}{n}$ or

$$\{x_n\}_{n=1}^{\infty} = \Big\{-1+\frac{1}{1}, 1+\frac{1}{2}, -1+\frac{1}{3}, 1+\frac{1}{4}, -1+\frac{1}{5}, 1+\frac{1}{6}, \cdots\Big\}.$$

Then for each $k \in \mathbb{N}$,

$$\sup_{n \geqslant k} x_n = 1 + \frac{1}{2[(k+1)/2]} \quad \text{and} \quad \inf_{n \geqslant k} x_n = -1 \,.$$

Therefore, $\limsup\limits_{n\to\infty} x_n = 1$ and $\liminf\limits_{n\to\infty} x_n = -1$.

Proposition 1.97. *Let $\{x_n\}_{n=1}^\infty$ be a sequence in $\mathbb{R}$. Then*

$$\limsup_{n\to\infty} -x_n = -\liminf_{n\to\infty} x_n \quad \textit{and} \quad \liminf_{n\to\infty} -x_n = -\limsup_{n\to\infty} x_n \,.$$

Proof. By the fact that $\sup\limits_{n \geqslant k} -x_n = -\inf\limits_{n \geqslant k} x_n$,

$$\limsup_{n\to\infty} -x_n = \lim_{k\to\infty} \sup_{n \geqslant k}(-x_n) = \lim_{k\to\infty} \big(-\inf_{n \geqslant k} x_n \big) = -\lim_{k\to\infty} \inf_{n \geqslant k} x_n = -\liminf_{n\to\infty} x_n \,.$$

The second identity holds simply by replacing x_n by $-x_n$ in the first identity. □

Proposition 1.98. *Let $\{x_n\}_{n=1}^\infty$ be a sequence in $\mathbb{R}$. Then*

1. $a = \liminf\limits_{n\to\infty} x_n \in \mathbb{R}$ *if and only if the following two statements hold*
   (a) *for all $\varepsilon > 0$, there exists $N > 0$ such that $a - \varepsilon < x_n$ whenever $n \geqslant N$; that is,*
   $$\forall\, \varepsilon > 0,\ \#\big\{n \in \mathbb{N} \,\big|\, x_n \leqslant a - \varepsilon\big\} < \infty;$$
   (b) *for all $\varepsilon > 0$ and $N > 0$, there exists $n \geqslant N$ such that $x_n < a + \varepsilon$; that is,*
   $$\forall\, \varepsilon > 0,\ \#\big\{n \in \mathbb{N} \,\big|\, x_n < a + \varepsilon\big\} = \infty \,.$$
2. $b = \limsup\limits_{n\to\infty} x_n \in \mathbb{R}$ *if and only if the following two statements hold*
   (a) *for all $\varepsilon > 0$, there exists $N > 0$ such that $b + \varepsilon > x_n$ whenever $n \geqslant N$; that is,*
   $$\forall\, \varepsilon > 0,\ \#\big\{n \in \mathbb{N} \,\big|\, x_n \geqslant b + \varepsilon\big\} < \infty;$$
   (b) *for all $\varepsilon > 0$ and $N > 0$, there exists $n \geqslant N$ such that $x_n > b - \varepsilon$; that is,*
   $$\forall\, \varepsilon > 0,\ \#\big\{n \in \mathbb{N} \,\big|\, x_n > b - \varepsilon\big\} = \infty \,.$$

Proof. We only prove 1 since the proof of 2 is similar. Let $z_k = \inf\limits_{n \geqslant k} x_n$, and

$$\sup_{k \geqslant 1} z_k = \lim_{k\to\infty} z_k = a \in \mathbb{R}^* \,.$$

We show that $a \in \mathbb{R}$ if and only if 1-(a) and 1-(b) both hold. Nevertheless, by Proposition 1.53 (or Remark 1.89), $a \in \mathbb{R}$ if and only if

(i) a is an upper bound for $\{z_k\}_{k=1}^{\infty}$.

(ii) $\forall\, \varepsilon > 0$, $\exists\, N \in \mathbb{N} \ni z_N > a - \varepsilon$.

We justify the equivalency between 1-(a) and (ii), as well as the equivalency between 1-(b) and (i) as follows:

(i) a is an upper bound for $\{z_k\}_{k=1}^{\infty} \Leftrightarrow a \geqslant z_k$ for all $k \in \mathbb{N} \Leftrightarrow \forall\, \varepsilon > 0, a + \varepsilon > z_k$ for all $k \in \mathbb{N} \Leftrightarrow \forall\, \varepsilon > 0$ and $k \in \mathbb{N}$, $a + \varepsilon > \inf\limits_{n \geqslant k} x_n \Leftrightarrow \forall\, \varepsilon > 0$ and $k \in \mathbb{N}$, $a + \varepsilon$ is not a lower bound for $\{x_n\}_{n \geqslant k}^{\infty} \Leftrightarrow \forall\, \varepsilon > 0$ and $k \in \mathbb{N}$, $\exists\, n \geqslant k \ni a + \varepsilon > x_n \Leftrightarrow$ 1-(b).

(ii) $\forall\, \varepsilon > 0$, $\exists\, N \in \mathbb{N} \ni z_N > a - \varepsilon \Leftrightarrow \forall\, \varepsilon > 0$, $\exists\, N > 0 \ni \inf\limits_{n \geqslant N} x_n > a - \varepsilon \Leftrightarrow \forall\, \varepsilon > 0$, $\exists\, N > 0$ such that $a - \varepsilon$ is a lower bound for $\{x_N, x_{N+1}, \cdots\} \Leftrightarrow \forall\, \varepsilon > 0$, $\exists\, N > 0$ such that $a - \varepsilon \leqslant x_n$ for all $n \geqslant N \Leftrightarrow \forall\, \varepsilon > 0$, $\exists\, N > 0$ such that $a - \varepsilon < x_n$ for all $n \geqslant N$ $\Leftrightarrow$ 1-(a). $\square$

Remark 1.99. By Proposition 1.98, if $a = \liminf\limits_{n\to\infty} x_n \in \mathbb{R}$, then

$$\forall\, \varepsilon > 0,\ \#\big\{n \in \mathbb{N} \,\big|\, x_n \in (a - \varepsilon, a + \varepsilon)\big\} = \infty$$

which implies that a is a cluster point of $\{x_n\}_{n=1}^{\infty}$. Moreover, 1-(a) of Proposition 1.98 implies that no other cluster points can be smaller than a. In other words, if $a = \liminf\limits_{n\to\infty} x_n \in \mathbb{R}$, then a is the smallest cluster point of $\{x_n\}_{n=1}^{\infty}$. Similarly, b is the largest cluster point of $\{x_n\}_{n=1}^{\infty}$ if $b = \limsup\limits_{n\to\infty} x_n \in \mathbb{R}$.

Theorem 1.100. *Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{R}$. Then*

1. $\liminf\limits_{n\to\infty} x_n \leqslant \limsup\limits_{n\to\infty} x_n$.
2. *If $\{x_n\}_{n=1}^{\infty}$ is bounded from above by M, then $\limsup\limits_{n\to\infty} x_n \leqslant M$.*
3. *If $\{x_n\}_{n=1}^{\infty}$ is bounded from below by m, then $\liminf\limits_{n\to\infty} x_n \geqslant m$.*
4. $\limsup\limits_{n\to\infty} x_n = \infty$ *if and only if $\{x_n\}_{n=1}^{\infty}$ is not bounded from above.*
5. $\liminf\limits_{n\to\infty} x_n = -\infty$ *if and only if $\{x_n\}_{n=1}^{\infty}$ is not bounded from below.*
6. *If x is a cluster point of $\{x_n\}_{n=1}^{\infty}$, then $\liminf\limits_{n\to\infty} x_n \leqslant x \leqslant \limsup\limits_{n\to\infty} x_n$.*
7. *If $a = \liminf\limits_{n\to\infty} x_n$ is finite, then a is a cluster point.*

8. *If* $b = \limsup_{n\to\infty} x_n$ *is finite, then* b *is a cluster point.*

9. *If* $\{x_n\}_{n=1}^{\infty}$ *converges to* x *in* $\mathbb{R}$ *if and only if* $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = x \in \mathbb{R}$.

Proof. Left as an exercise. □

Remark 1.101. Using the definition of cluster points of a sequence in Remark 1.89, Remark 1.99 and Theorem 1.100 together imply that the limit superior/inferior of a sequence is the largest/smallest cluster point of that sequence.

Example 1.102. Let $\mathbb{Q} \cap [0,1] = \{q_1, q_2, \cdots, q_n, \cdots\}$. Then $\{q_n\}_{n=1}^{\infty}$ does not converge since $\limsup_{n\to\infty} q_n = 1$ while $\liminf_{n\to\infty} q_n = 0$ by Example 1.65.

1.9 Exercises

§ 1.1 Ordered Fields

Problem 1.1. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, $x, y \in \mathbb{F}$, and $n \in \mathbb{N}$. Show that

1. If $0 \leqslant x < y$, then $x^n < y^n$.

2. If $0 \leqslant x, y$ and $x^n < y^n$, then $x < y$.

Problem 1.2. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, and $a, b \in \mathbb{F}$. Show that $a \leqslant b$ if and only if for all $\varepsilon > 0$, $a < b + \varepsilon$.

Problem 1.3. Prove Proposition 1.17.

§ 1.2 Sequences in Ordered Fields

Problem 1.4. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an Archimedean ordered field, and $x, y \in \mathbb{F}$ satisfying $x < y$. Show that there exists $r \in \mathbb{Q}$ such that $x < r < y$. This property is called the ***denseness*** of $\mathbb{Q}$ (in Archimedean ordered fields).

Problem 1.5. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an Archimedean ordered field, and $0 < \alpha < 1$. Show that $\lim_{n\to\infty} \alpha^n = 0$.

Problem 1.6. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field with Archimedean Property, $I \subseteq \mathbb{F}$ be a non-empty interval, and $f : I \to \mathbb{F}$ be a function.

1. f is said to have a limit at $c \in I$ or we say that the limit of f at c exists if

$$\lim_{n\to\infty} f(x_n) \text{ exists for all convergent sequences } \{x_n\}_{n=1}^\infty \subseteq I\backslash\{c\} \text{ with limit } c.$$

Show that the limit of f at c exists if and only if there exists $L \in \mathbb{F}$ satisfying that for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\big|f(x) - L\big| < \varepsilon \qquad \text{whenever} \quad 0 < |x - c| < \delta \text{ and } x \in I\,.$$

2. f is said to be continuous at a point $c \in I$ if

$$\lim_{n\to\infty} f(x_n) = f(c) \quad \text{for all convergent sequences } \{x_n\}_{n=1}^\infty \subseteq I \text{ with limit } c\,.$$

Show that f is continuous at c if and only if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\big|f(x) - f(c)\big| < \varepsilon \qquad \text{whenever} \quad |x - c| < \delta \text{ and } x \in I\,.$$

§ 1.3 Monotone Sequence Property

Problem 1.7. Complete the following.

1. Verify the Wallis's formula: if n is a non-negative integer, then

$$\int_0^{\frac{\pi}{2}} \sin^{2n+1} x\, dx = \int_0^{\frac{\pi}{2}} \cos^{2n+1} x\, dx = \frac{(2^n n!)^2}{(2n+1)!}$$

and

$$\int_0^{\frac{\pi}{2}} \sin^{2n} x\, dx = \int_0^{\frac{\pi}{2}} \cos^{2n} x\, dx = \frac{(2n)!}{(2^n n!)^2} \cdot \frac{\pi}{2}\,.$$

2. Let $I_n = \displaystyle\int_0^{\frac{\pi}{2}} \sin^n x\, dx$. Show that $\displaystyle\lim_{n\to\infty} \frac{I_{2n+1}}{I_{2n}} = 1$.

3. Let $s_n = \dfrac{n!}{n^{n+0.5}e^{-n}}$. Show that $\{s_n\}_{n=1}^\infty$ is a decreasing sequence.

4. Suppose that you know that $\mathbb{R}$ satisfies the monotone sequence property. Show that $\{s_n\}_{n=1}^\infty$ converges and find the limit of $\{s_n\}_{n=1}^\infty$.

Hint: 2. Show that $I_{2n+2} \leqslant I_{2n+1} \leqslant I_{2n}$ for all $n \in \mathbb{N}$ and then apply the Sandwich lemma.

3. Consider the function $f(x) = \Big(1 + \dfrac{1}{x}\Big)^{x+0.5}$.

Problem 1.8. Let $(\mathbb{F}, +, \cdot, \leqslant)$ an ordered field satisfying the monotone sequence property, and $y \in \mathbb{F}$ satisfying $y > 1$. Complete the following.

1. Define $y^{1/n}$ properly. (Hint: see how we define $\sqrt{y}$ in Example 1.48).

2. Show that $y^n - 1 > n(y-1)$ for all $n \in \mathbb{N}\backslash\{1\}$; thus $y - 1 > n(y^{1/n} - 1)$.

3. Show that if $t > 1$ and $n > (y-1)/(t-1)$, then $y^{1/n} < t$.

4. Show that $\lim\limits_{n\to\infty} y^{1/n} = 1$ as $n \to \infty$.

§ 1.4 Least Upper Bound Property

Problem 1.9. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an Archimedean ordered field, and $S = (0,1)$. Show that $\inf S = 0$ and $\sup S = 1$.

Problem 1.10. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field satisfying the least upper bound property, and A be a non-empty set of $\mathbb{F}$ which is bounded from below. Define the set $-A$ by $-A \equiv \left\{ -x \in \mathbb{F} \,\middle|\, x \in A \right\}$. Prove that

$$\inf A = -\sup(-A)\,.$$

Note that this problem shows that if $\mathbb{F}$ satisfies the least upper bound property, then $\mathbb{F}$ satisfies the greatest lower bound property.

Problem 1.11. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field satisfying the least upper bound property, and A be a non-empty set of $\mathbb{F}$ which is bounded from above. Define the set A^{-1} by $A^{-1} \equiv \left\{ x^{-1} \in \mathbb{F} \,\middle|\, x \in A \right\}$ whenever $0 \notin A$. Prove that

$$\inf A^{-1} = \frac{1}{\sup A} \qquad \text{whenever } A \text{ consists of only positive elements in } \mathbb{F}\,.$$

Is $\inf A^{-1} = \dfrac{1}{\sup A}$ in general true?

Problem 1.12. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field (satisfying the least upper bound property), and A be a non-empty subset of $\mathbb{F}$. Show that if $a \in A$ is an upper bound for A, then a is the least upper bound for A.

Problem 1.13. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field satisfying the least upper bound property, and A, B be non-empty subsets of $\mathbb{F}$. Define $A + B = \{x + y \,|\, x \in A, y \in B\}$. Justify if the following statements are true or false by providing a proof for the true statement and giving a counter-example for the false ones.

1. $\sup(A+B) = \sup A + \sup B$.
2. $\inf(A+B) = \inf A + \inf B$.
3. $\sup(A \cap B) \leqslant \min\{\sup A, \sup B\}$.
4. $\sup(A \cap B) = \min\{\sup A, \sup B\}$.
5. $\sup(A \cup B) \geqslant \max\{\sup A, \sup B\}$.
6. $\sup(A \cup B) = \max\{\sup A, \sup B\}$.

Problem 1.14. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field satisfying the least upper bound property, and $S \subseteq \mathbb{F}$ be non-empty.

1. Show that if S is bounded from below, then
$$\inf S = \sup \big\{x \in \mathbb{F} \,\big|\, x \text{ is a lower bound for } S\big\}$$
2. Show that if S is bounded from above, then
$$\sup S = \inf \big\{x \in \mathbb{F} \,\big|\, x \text{ is an upper bound for } S\big\}.$$

Problem 1.15. Let A, B be two sets, and $f : A \times B \to \mathbb{F}$ be a function, where $(\mathbb{F}, +, \cdot, \leqslant)$ is an ordered field satisfying the least upper bound property. Show that

$$\sup_{(x,y)\in A\times B} f(x,y) = \sup_{y\in B}\big(\sup_{x\in A} f(x,y)\big) = \sup_{x\in A}\big(\sup_{y\in B} f(x,y)\big)$$

and

$$\inf_{(x,y)\in A\times B} f(x,y) = \inf_{y\in B}\big(\inf_{x\in A} f(x,y)\big) = \inf_{x\in A}\big(\inf_{y\in B} f(x,y)\big).$$

Problem 1.16. Let A be a set, and $f, g : A \to \mathbb{R}$ be two functions. Let $h = \max\{f, g\}$; that is,

$$h(x) = \max\big\{f(x), g(x)\big\} \qquad \forall\, x \in A.$$

Show that

$$\sup_{x\in A} h(x) = \max\Big\{\sup_{x\in A} f(x), \sup_{x\in A} g(x)\Big\}.$$

Generalize the result above to the following: if $f_1, \cdots, f_n : A \to \mathbb{R}$ are real-valued functions, then

$$\sup_{x\in A} \max\big\{f_1(x), \cdots, f_n(x)\big\} = \max\Big\{\sup_{x\in A} f_1(x), \sup_{x\in A} f_2(x), \cdots, \sup_{x\in A} f_n(x)\Big\}.$$

Can one conclude that if $f_n : A \to \mathbb{R}$ is a sequence of functions, then

$$\sup_{x\in A} \sup\big\{f_1(x), \cdots, f_n(x), \cdots\big\} = \sup\Big\{\sup_{x\in A} f_1(x), \sup_{x\in A} f_2(x), \cdots, \sup_{x\in A} f_n(x), \cdots\Big\}.$$

§ 1.5 Bolzano-Weierstrass Property

Problem 1.17. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an ordered field, and $\{x_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{F}$. Show that the following three statements are equivalent.

1. $\{x_n\}_{n=1}^{\infty}$ converges.

2. Every proper subsequence of $\{x_n\}_{n=1}^{\infty}$ converges.

3. Every subsequence of $\{x_n\}_{n=1}^{\infty}$ converges.

Problem 1.18. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an Archimedean ordered field, $\{x_n\}_{n=1}^{\infty}$ be a sequence in $\mathbb{F}$, and $x \in \mathbb{F}$. Show the following.

1. x is a cluster point of $\{x_n\}_{n=1}^{\infty}$ if and only if for all $\varepsilon > 0$ and $N > 0$ there exists $n \geqslant N$ such that $|x_n - x| < \varepsilon$.

2. x is a cluster point of $\{x_n\}_{n=1}^{\infty}$ if and only if there exists a subsequence $\{x_{n_j}\}_{j=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ converges to x.

Remark: The Archimedean Property is only used in 2 in the direction "$\Rightarrow$".

Problem 1.19. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an Archimedean ordered field. A number $x \in \mathbb{F}$ is called an ***accumulation point*** of a set $A \subseteq \mathbb{F}$ if for all $\delta > 0$, $(x - \delta, x + \delta)$ contains at least one point of A distinct from x. In logic notation,

$$x \text{ is an accumulation point of } A \quad \Leftrightarrow \quad (\forall\, \delta > 0)\big(A \cap (x - \delta, x + \delta)\backslash\{x\} \neq \varnothing\big)\,.$$

1. Show that if $\{x_n\}_{n=1}^{\infty}$ is a sequence in $\mathbb{F}$ so that $x_i \neq x_j$ for all $i, j \in \mathbb{N}$ and $A = \big\{x_k \,\big|\, k \in \mathbb{N}\big\}$, then x is an accumulation of A if and only if x is a cluster point of $\{x_n\}_{n=1}^{\infty}$.

2. How about if the condition $x_i \neq x_j$ for all $i, j \in \mathbb{N}$ is removed? Is the statement in 1 still valid?

§ 1.6 Cauchy Sequences

Problem 1.20. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an Archimedean ordered field, and $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{F}$ be a sequence satisfying $|x_n - x_{n+1}| < \dfrac{1}{n}$ for all $n \in \mathbb{N}$. Prove or disprove that there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ so that $\{x_{n_k}\}_{k=1}^{\infty}$ is a Cauchy sequence.

Problem 1.21. Let $(\mathbb{F}, +, \cdot, \leqslant)$ be an Archimedean ordered field, and $f : \mathbb{F} \to \mathbb{F}$ be a function so that

$$|f(x) - f(y)| \leqslant \alpha |x - y| \qquad \forall\, x, y \in \mathbb{F}\,,$$

where $\alpha \in \mathbb{F}$ is a constant satisfying $0 < \alpha < 1$. Pick an arbitrary $x_1 \in \mathbb{F}$, and define $x_{k+1} = f(x_k)$ for all $k \in \mathbb{N}$. Show that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{F}$.

§ 1.7 Completeness

Problem 1.22. Let $a \in \mathbb{R}$. Define a_n through the iterated relation

$$a_n = a_{n-1}^2 - a_{n-1} + 1 \qquad \forall\, n > 1, a_1 = a\,.$$

For what a is the sequence $\{a_n\}_{n=1}^{\infty}$ (1) monotone? (2) bounded? (3) convergent? Compute the limit in the case of convergence.

Problem 1.23. Suppose that $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ are two Cauchy sequences in $\mathbb{R}$. Show that the sequence $\big\{|x_n - y_n|\big\}_{n=1}^{\infty}$ converges.

Problem 1.24. Let $b \in \mathbb{R}$, $r \in \mathbb{R}$, and $b > 1$, $r > 0$. Show that

$$\sup_{x \in (1,b)} x^r = \Big(\sup_{x \in (1,b)} x\Big)^r = b^r\,.$$

Problem 1.25. Let $b \in \mathbb{R}$ and $b > 1$.

1. Show the law of exponents holds (for rational exponents); that is, show that

   (a) if r, s in $\mathbb{Q}$, then $b^{r+s} = b^r \cdot b^s$.

   (b) if r, s in $\mathbb{Q}$, then $b^{r \cdot s} = (b^r)^s$.

2. For $x \in \mathbb{F}$, let $B(x) = \big\{b^t \in \mathbb{F} \,\big|\, t \in \mathbb{Q}, t \leqslant x\big\}$. Show that $\sup B(x)$ exists for all $x \in \mathbb{F}$, and $b^r = \sup B(r)$ if $r \in \mathbb{Q}$.

3. Define $b^x = \sup B(x)$ for $x \in \mathbb{F}$. Show that $B(x) > 0$ for all $x \in \mathbb{F}$ and the law of exponents (for exponents in $\mathbb{F}$)

   (a) if x, y in $\mathbb{F}$, then $b^{x+y} = b^x \cdot b^y\,,$ (b) if $x, y > 0$, then $b^{x \cdot y} = (b^x)^y\,,$

   are also valid.

4. Show that if $x_1, x_2 \in \mathbb{F}$ and $x_1 < x_2$, then $b^{x_1} < b^{x_2}$. This implies that if x_1, x_2 are two numbers in $\mathbb{F}$ satisfying $b^{x_1} = b^{x_2}$, then $x_1 = x_2$.

5. Let $y > 0$ be given. Show that if $u, v \in \mathbb{F}$ such that $b^u < y$ and $b^v > y$, then $b^{u+1/n} < y$ and $b^{v-1/n} > y$ for sufficiently large n.

6. Let $y > 0$ be given, and $A \subseteq \mathbb{F}$ be the set of all w such that $b^w < y$. Show that $\sup A$ exists and $x = \sup A$ satisfies $b^x = y$. The number x (the uniqueness is guaranteed by 4) satisfying $b^x = y$ is called the logarithm of y to the base b, and is denoted by $\log_b y$.

Hint: Make use of Problem 1.8.

Problem 1.26. In this problem we prove the Intermediate Value Theorem:

> Let $f : [a, b] \to \mathbb{R}$ be continuous (at every point of $[a, b]$); that is,
>
> $$\lim_{n\to\infty} f(x_n) = f\big(\lim_{n\to\infty} x_n\big) \quad \text{for all convergent sequence } \{x_n\}_{n=1}^{\infty} \subseteq [a, b].$$
>
> If $f(a)f(b) < 0$, then there exists $c \in [a, b]$ such that $f(c) = 0$.

Complete the following.

1. W.L.O.G, we can assume that $f(a) < 0$. Define the set $S = \{x \in [a, b] \,|\, f(x) > 0\}$. Show that $\inf S$ exists.

2. Let $c = \inf S$. Show that $f(c) \geqslant 0$.

3. Conclude that $f(c) \leqslant 0$ as well.

Hint:

1. For 2, show that there exists a sequence $\{c_n\}_{n=1}^{\infty}$ in S such that $c_n \to c$ as $n \to \infty$.

2. For 3, show that there exists a sequence $\{c_n\}_{n=1}^{\infty}$ in $[a, c)$ such that $c_n \to c$ as $n \to \infty$.

Problem 1.27. In this problem we prove the Extreme Value Theorem:

> Let $f : [a, b] \to \mathbb{R}$ be continuous (at every point of $[a, b]$); that is,
>
> $$\lim_{n\to\infty} f(x_n) = f\big(\lim_{n\to\infty} x_n\big) \quad \text{for all convergent sequence } \{x_n\}_{n=1}^{\infty} \subseteq [a, b].$$
>
> Then there exist $c, d \in [a, b]$ such that $f(c) = \sup\limits_{x\in[a,b]} f(x)$ and $f(d) = \inf\limits_{x\in[a,b]} f(x)$.

Complete the following.

1. Show that there exist sequences $\{c_n\}_{n=1}^\infty$ and $\{d_n\}_{n=1}^\infty$ in $[a,b]$ such that
$$\lim_{n\to\infty} f(c_n) = \sup_{x\in[a,b]} f(x) \qquad \text{and} \qquad \lim_{n\to\infty} f(d_n) = \inf_{x\in[a,b]} f(x)\,.$$

2. Apply the Bolzano-Weierstrass Property of $\mathbb{R}$ to extract convergent subsequences $\{c_{n_k}\}_{k=1}^\infty$ and $\{d_{n_k}\}_{k=1}^\infty$ with limit c and d, respectively. Show that $c, d \in [a,b]$.

3. Show that $f(c) = \sup\limits_{x\in[a,b]} f(x)$ and $f(d) = \inf\limits_{x\in[a,b]} f(x)$.

§ 1.8 Limit Inferior, Limit Superior

Problem 1.28. Let $\{x_n\}_{n=1}^\infty$ and $\{y_n\}_{n=1}^\infty$ be sequences in $\mathbb{R}$. Prove the following inequalities:

$$\begin{aligned}
\liminf_{n\to\infty} x_n + \liminf_{n\to\infty} y_n &\leqslant \liminf_{n\to\infty}(x_n+y_n) \leqslant \liminf_{n\to\infty} x_n + \limsup_{n\to\infty} y_n \\
&\leqslant \limsup_{n\to\infty}(x_n+y_n) \leqslant \limsup_{n\to\infty} x_n + \limsup_{n\to\infty} y_n\,; \\
\big(\liminf_{n\to\infty}|x_n|\big)\big(\liminf_{n\to\infty}|y_n|\big) &\leqslant \liminf_{n\to\infty}|x_n y_n| \leqslant \big(\liminf_{n\to\infty}|x_n|\big)\big(\limsup_{n\to\infty}|y_n|\big) \\
&\leqslant \limsup_{n\to\infty}|x_n y_n| \leqslant \big(\limsup_{n\to\infty}|x_n|\big)\big(\limsup_{n\to\infty}|y_n|\big)\,.
\end{aligned}$$

Give examples showing that the equalities are generally not true.

Problem 1.29. Prove that

$$\liminf_{n\to\infty}\frac{|x_{n+1}|}{|x_n|} \leqslant \liminf_{n\to\infty}\sqrt[n]{|x_n|} \leqslant \limsup_{n\to\infty}\sqrt[n]{|x_n|} \leqslant \limsup_{n\to\infty}\frac{|x_{n+1}|}{|x_n|}\,.$$

Give examples to show that the equalities are not true in general. Is it true that $\lim\limits_{n\to\infty}\sqrt[n]{|x_n|}$ exists implies that $\lim\limits_{n\to\infty}\dfrac{|x_{n+1}|}{|x_n|}$ also exists?

Problem 1.30. Find the following limits.

$$\lim_{n\to\infty}\frac{1}{n}\sqrt[n]{n!}\,, \qquad \lim_{n\to\infty}\frac{1}{n}\sqrt[n]{(n+1)(n+2)\cdots(2n)}\,.$$

Problem 1.31. Given the following sets consisting of elements of some sequence of real numbers. Find the limsup and liminf of the sequence.

1. $\big\{\cos m \,\big|\, m = 0, 1, 2, \cdots\big\}$.

2. $\left\{ \sqrt[m]{|\sin m|} \,\middle|\, m = 1, 2, \cdots \right\}$.

3. $\left\{ \left(1 + \frac{1}{m}\right) \sin \frac{m\pi}{6} \,\middle|\, m = 1, 2, \cdots \right\}$.

Hint: 1. First show that for all irrational α, the set

$$S = \left\{ x \in [0, 1] \,\middle|\, x = k\alpha \pmod 1 \text{ for some } k \in \mathbb{N} \right\}$$

is dense in $[0, 1]$; that is, for all $y \in [0, 1]$ and $\varepsilon > 0$, there exists $x \in S \cap (y - \varepsilon, y + \varepsilon)$. Then choose $\alpha = \frac{1}{2\pi}$ to conclude that

$$T = \left\{ x \in [0, 2\pi] \,\middle|\, x = k \pmod{2\pi} \text{ for some } k \in \mathbb{N} \right\}$$

is dense in $[0, 2\pi]$. To prove that S is dense in $[0, 1]$, you might want to consider the following set

$$S_k = \left\{ x \in [0, 1] \,\middle|\, x = \ell\alpha \pmod 1 \text{ for some } 1 \leqslant \ell \leqslant k + 1 \right\}$$

Note that there must be two points in S_k whose distance is less than $\frac{1}{k}$. What happened to (the multiples of) the difference of these two points?
2. Use the fact that π is a Liouville number; that is, there exists $d \in \mathbb{N}$ such that

$$\left| \pi - \frac{p}{q} \right| \geqslant \frac{1}{q^d} \qquad \forall\, p, q \in \mathbb{Z}, q \neq 0\,.$$

§ 1.9 Chapter Review

Problem 1.32 (True or False). Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. Let $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ be a sequence and $\limsup\limits_{n\to\infty} x_n = x$. Then $\sup\limits_{n\in\mathbb{N}} x_n = x$.

2. Let $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ be a sequence such that $|x_n - x_{n+1}| \leqslant \frac{1}{n}$. Then $\{x_n\}_{n=1}^{\infty}$ converges in $\mathbb{R}$.

3. If a bounded sequence $\{x_n\}_{n=1}^{\infty}$ in $\mathbb{R}$ satisfies $x_{n+1} - \epsilon_n \leqslant x_n$ for $n \in \mathbb{N}$, where $\sum\limits_{n=1}^{\infty} \epsilon_n$ is an absolute convergent series; that is, the partial sum $\sum\limits_{n=1}^{k} |\epsilon_n|$ converges as $k \to \infty$, then $\{x_n\}_{n=1}^{\infty}$ converges.

4. Let $\pi : \mathbb{N} \to \mathbb{N}$ be one-to-one and onto (such map π is called a rearrangement), and $\{x_n\}_{n=1}^{\infty}$ is a convergent sequence. Then $\{x_{\pi(n)}\}_{n=1}^{\infty}$ is also convergent.

5. Any rearrangement of the series $\sum\limits_{n=1}^{\infty} x_n$ diverges if and only if x_n does not tend to 0 as $n \to \infty$.

6. If $\{x_n\}_{n=1}^{\infty}$ is a sequence of distinct non-zero real numbers such that $\lim\limits_{n\to\infty} x_n = 0$, then the set $\{mx_n \,|\, m \in \mathbb{Z}, n \in \mathbb{N}\}$ is dense in $\mathbb{R}$.

Chapter 2

Normed Vector Spaces and Metric Spaces

2.1 Euclidean Spaces and Vector Spaces

Definition 2.1. ***Euclidean*** $\boldsymbol{n}$***-space***, denoted by $\mathbb{R}^n$, consists of all ordered n-tuples of real numbers. Symbolically,

$$\mathbb{R}^n = \left\{\boldsymbol{x} \,\middle|\, \boldsymbol{x} = (x_1, x_2, \cdots, x_n), x_i \in \mathbb{R}\right\}.$$

Elements of $\mathbb{R}^n$ are generally denoted by single letters that stand for n-tuples such as $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, and speak of $\boldsymbol{x}$ as a "point" in $\mathbb{R}^n$.

Remark 2.2. Let $\mathbb{C}$ denote the collection of ordered pairs $\mathbb{C} = \left\{(a, b) \,\middle|\, a, b \in \mathbb{R}\right\}$ on which $+$ and $\cdot$ are given by Example 1.6. Then $\mathbb{C}$ is a field, and is called the complex number system. The ordered pair (a, b) in $\mathbb{C}$ is usually denoted by $a + bi$, where $i^2 = -1$ according to the definition of the multiplication. The space $\mathbb{C}^n$ can be defined similarly by

$$\mathbb{C}^n = \left\{\boldsymbol{z} \,\middle|\, \boldsymbol{z} = (z_1, z_2, \cdots, z_n), z_i \in \mathbb{C}\right\}.$$

Definition 2.3. A ***vector space*** $\mathcal{V}$ over a scalar field $\mathbb{F}$ is a set of elements called vectors, with given operations of vector addition $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ and scalar multiplication $\cdot : \mathbb{F} \times \mathcal{V} \to \mathcal{V}$ such that

1. $\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{v}$ for all $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

2. $(\boldsymbol{v} + \boldsymbol{w}) + \boldsymbol{u} = \boldsymbol{v} + (\boldsymbol{u} + \boldsymbol{w})$ for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

3. there exists $\mathbf{0}$, the zero vector, such that $\boldsymbol{v} + \mathbf{0} = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathcal{V}$.

4. for each $\boldsymbol{v} \in \mathcal{V}$ there exists $\boldsymbol{w} \in \mathcal{V}$ such that $\boldsymbol{v} + \boldsymbol{w} = \mathbf{0}$.

5. $\lambda \cdot (\boldsymbol{v} + \boldsymbol{w}) = \lambda \cdot \boldsymbol{v} + \lambda \cdot \boldsymbol{w}$ for all $\lambda \in \mathbb{F}$ and $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

6. $(\lambda + \mu) \cdot \boldsymbol{v} = \lambda \cdot \boldsymbol{v} + \mu \cdot \boldsymbol{v}$ for all $\lambda, \mu \in \mathbb{F}$ and $\boldsymbol{v} \in \mathcal{V}$.

7. $(\lambda \cdot \mu) \cdot \boldsymbol{v} = \lambda \cdot (\mu \cdot \boldsymbol{v})$ for all $\lambda, \mu \in \mathbb{F}$ and $\boldsymbol{v} \in \mathcal{V}$.

8. $1 \cdot \boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathcal{V}$.

Remark 2.4. In general the scalar field $\mathbb{F}$ can be the rational number system $\mathbb{Q}$, the real number system $\mathbb{R}$, or even the complex number system $\mathbb{C}$. In this lecture note, $\mathbb{F}$ is taken as either the real number system $\mathbb{R}$ or the complex number system $\mathbb{C}$ (and mostly $\mathbb{R}$ if not specified).

Example 2.5. Let the vector addition and scalar multiplication on $\mathbb{F}^n$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, be defined by

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, \cdots, x_n + y_n) \quad \text{if} \quad \boldsymbol{x} = (x_1, \cdots, x_n), \boldsymbol{y} = (y_1, \cdots, y_n)$$

and

$$\lambda \cdot \boldsymbol{x} = (\lambda x_1, \cdots, \lambda x_n) \quad \text{if} \quad \lambda \in \mathbb{F}, \boldsymbol{x} = (x_1, \cdots, x_n) \in \mathbb{F}^n .$$

Then $\mathbb{F}^n$ is a vector space over $\mathbb{F}$ if $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Moreover, $\mathbb{C}^n$ is a vector space over $\mathbb{R}$; however, $\mathbb{R}^n$ is not a vector space over $\mathbb{C}$.

Example 2.6. Let $\mathcal{M}_{n\times m}$ be the collection of all $n \times m$ real matrices; that is, $\mathcal{M}_{n\times m} \equiv \{n \times m \text{ matrix with entries in } \mathbb{R}\}$. Define

$$A + B \equiv [a_{ij} + b_{ij}], \quad \lambda \cdot A \equiv [\lambda \cdot a_{ij}] \quad \text{if} \quad \lambda \in \mathbb{R}, A = [a_{ij}], B = [b_{ij}] \in \mathcal{M} .$$

Then $\mathcal{M}_{n\times m}$ is a vector space over $\mathbb{R}$.

Definition 2.7. $\mathcal{W}$ is called a ***subspace*** of a vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$ if

1. $\mathcal{W}$ is a subset of $\mathcal{V}$.

2. $(\mathcal{W}, +, \cdot)$, with vector addition and scalar multiplication in $\mathcal{V}$, is a vector space over $\mathbb{F}$.

Example 2.8. $\mathcal{V} = \mathbb{R}^3$, $W = \mathbb{R}^2 \times \{0\} \equiv \{(x, y, 0) \,|\, x, y \in \mathbb{R}\}$. $\mathcal{W}$ is a subspace of $\mathcal{V}$.

Lemma 2.9. *If $\mathcal{W}$ is a subset of a vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$, then $\mathcal{W}$ is a subspace if and only if $\lambda \cdot \boldsymbol{v} + \mu \cdot \boldsymbol{w} \in \mathcal{W}$ for all $\lambda, \mu \in \mathbb{F}$, $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{W}$.*

Remark 2.10. "n" is called the ***dimension*** of $\mathbb{R}^n$.
There are n linearly independent vectors $\mathbf{e}_1 = (1, 0, \cdots, 0), \mathbf{e}_2 = (0, 1, 0, \cdots, 0), \cdots, \mathbf{e}_n = (0, 0, \cdots, 0, 1)$, but if $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{n+1}$ are $(n+1)$ vectors in $\mathbb{R}^n$, there exists $\lambda_1, \cdots, \lambda_{n+1} \in \mathbb{R}$ such that $(\lambda_1, \cdots, \lambda_{n+1}) \neq (0, \cdots, 0)$ and $\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_{n+1} \boldsymbol{v}_{n+1} = \mathbf{0}$.

On the other hand, the dimension of $\mathbb{C}^n$ depends on the scalar field $\mathbb{F}$.

1. If $\mathbb{F} = \mathbb{R}$, then the dimension of $\mathbb{C}^n$ is $2n$ since $\mathbf{e}_1, \cdots, \mathbf{e}_n, i\mathbf{e}_1, \cdots, i\mathbf{e}_n$ are $2n$ linearly independent vectors in $\mathbb{C}^n$, and any $(2n+1)$ non-zeros vectors in $\mathbb{C}^n$ are not linearly independent; thus the dimension of $\mathbb{C}^n$ over $\mathbb{R}$ is $2n$. When $\mathbb{F} = \mathbb{R}$, we usually identify $\mathbb{C}^n$ as $\mathbb{R}^{2n}$.

2. If $\mathbb{F} = \mathbb{C}$, $\mathbf{e}_1, \cdots, \mathbf{e}_n$ are linearly independent in $\mathbb{C}^n$ and any $(n+1)$ non-zeros vectors in $\mathbb{C}^n$ are not linearly independent; thus the dimension of $\mathbb{C}^n$ over $\mathbb{C}$ is n.

Definition 2.11. Let $\mathcal{V}$ be a vector space (over a scalar field $\mathbb{F}$), and A, B be subsets of $\mathcal{V}$. The sum of A and B, denoted by $A + B$, is the set $\{\boldsymbol{a} + \boldsymbol{b} \mid \boldsymbol{a} \in A, \boldsymbol{b} \in B\}$. If A consists of only one single vector $\boldsymbol{a}$, $A + B$ is usually denoted by $\boldsymbol{a} + B$ instead of $\{\boldsymbol{a}\} + B$.

The following theorem should be clear to the readers, and is left as an exercise.

Theorem 2.12. *Let $\mathcal{V}$ be a vector space (over a field $\mathbb{F}$), and A, B be subsets of $\mathcal{V}$. Then*

$$A + B = \bigcup_{\boldsymbol{a} \in A} (\boldsymbol{a} + B) = \bigcup_{\boldsymbol{b} \in B} (\boldsymbol{b} + A).$$

Definition 2.13. A subset $H \subseteq \mathbb{R}^n$ is called a ***hyperplane*** or ***hyperspace*** if H is $(n-1)$-dimensional subspace of $\mathbb{R}^n$. An ***affine hyperplane*** is a set $\boldsymbol{x} + H$ for some $\boldsymbol{x} \in \mathbb{R}^n$ and hyperplane H.

Example 2.14. A straight line on the plane is a hyperplane, and a plane on the (3-dimensional) space is also a hyperplane. However, a straight line on the (3-dimensional) space is not a hyperplane.

2.2 Normed Vector Spaces, Inner Product Spaces and Metric Spaces

Definition 2.15. A ***normed vector space*** (or simply ***normed space***) $(\mathcal{V}, \|\cdot\|)$ is a vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, associated with a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ such that

(a) $\|\boldsymbol{x}\| \geqslant 0$ for all $\boldsymbol{x} \in \mathcal{V}$.

(b) $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = \mathbf{0}$.

(c) $\|\lambda \cdot \boldsymbol{x}\| = |\lambda| \cdot \|\boldsymbol{x}\|$ for all $\lambda \in \mathbb{F}$ and $\boldsymbol{x} \in \mathcal{V}$.

(d) $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$.

A function $\|\cdot\|$ satisfying (a)-(d) is called a ***norm*** on $\mathcal{V}$.

Remark 2.16. Let $(\mathcal{V}, \|\cdot\|)$ be a normed vector space. Treating a vector in $\mathcal{V}$ as a point in $\mathcal{V}$, the number $\|\boldsymbol{x} - \boldsymbol{y}\|$ can be viewed as the distance (induced by the norm) between $\boldsymbol{x}$ and $\boldsymbol{y}$, and (d) implies that

$$\|\boldsymbol{x} - \boldsymbol{y}\| \leqslant \|\boldsymbol{x} - \boldsymbol{z}\| + \|\boldsymbol{z} - \boldsymbol{y}\| \qquad \forall\, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathcal{V}.$$

The inequality above states that the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ is not greater than the sum of the distance between $\boldsymbol{x}$ and $\boldsymbol{z}$ and the distance between $\boldsymbol{y}$ and $\boldsymbol{z}$; thus the inequality in (d) is called the ***triangle inequality***.

Example 2.17. Let $\mathcal{V} = \mathbb{R}^n$, and define

$$\|\boldsymbol{x}\|_p \equiv \begin{cases} \Big(\sum\limits_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} & \text{if } 1 \leqslant p < \infty, \\ \max\big\{|x_1|, \cdots, |x_n|\big\} & \text{if } p = \infty, \end{cases} \qquad \text{for all } \boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n.$$

Then $\|\cdot\|_p$ is a norm, called p-norm, on $\mathbb{R}^n$. Property (d) in Definition 2.15; that is, $\|\boldsymbol{x} + \boldsymbol{y}\|_p \leqslant \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$ (so-called the ***Minkowski inequality***), is left as an exercise.

Example 2.18. Let $\mathcal{V} = \mathbb{C}$ and the norm $\|\cdot\|$ is the usual absolute value of complex numbers; that is, $\|a + ib\| \equiv |a + ib| = \sqrt{a^2 + b^2}$. Then $(\mathbb{C}, \|\cdot\|)$ is a normed vector space.

Example 2.19. Let $\mathcal{M}_{n\times m} \equiv \big\{n \times m$ matrix with entries in $\mathbb{R}\big\}$, and $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_m\}$ be standard basis of $\mathbb{R}^m$. For $A = \big[a_{ij}\big] \in \mathcal{M}_{n\times m}$ and $\boldsymbol{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_m\mathbf{e}_m$, we use $A\boldsymbol{x}$ to denote the n-vector

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}.$$

Define

$$\|A\|_p = \sup_{\|\boldsymbol{x}\|_p=1} \|A\boldsymbol{x}\|_p = \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|A\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p} \qquad \forall\, A \in \mathcal{M}_{n\times m}\,;$$

that is, $\|A\|_p$ is the least upper bound of the set $\left\{\dfrac{\|A\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p} \,\middle|\, \boldsymbol{x}\neq\boldsymbol{0}, \boldsymbol{x}\in\mathbb{R}^m\right\}$. Then $\|\cdot\|_p$ satisfies the triangle inequality for the following reason: Suppose that $A, B \in \mathcal{M}_{n\times m}$. If $\|\boldsymbol{x}\|_p = 1$,

$$\begin{aligned}\|(A+B)\boldsymbol{x}\|_p &= \|A\boldsymbol{x}+B\boldsymbol{x}\|_p \leqslant \|A\boldsymbol{x}\|_p + \|B\boldsymbol{x}\|_p \\ &\leqslant \sup_{\|\boldsymbol{x}\|_p=1} \|A\boldsymbol{x}\|_p + \sup_{\|\boldsymbol{x}\|_p=1} \|B\boldsymbol{x}\|_p = \|A\|_p + \|B\|_p\,;\end{aligned}$$

thus

$$\|A+B\|_p = \sup_{\|\boldsymbol{x}\|_p=1} \|(A+B)\boldsymbol{x}\|_p \leqslant \|A\|_p + \|B\|_p\,.$$

Since property (a), (b), (c) in Definition 2.15 are obvious, we conclude that $\|\cdot\|_p$ is a norm on $\mathcal{M}_{n\times m}$. Moreover, by the definition of the p-norm we have $\dfrac{\|A\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p} \leqslant \|A\|_p$ for all $\boldsymbol{x}\neq\boldsymbol{0}$; thus

$$\|A\boldsymbol{x}\|_p \leqslant \|A\|_p\|\boldsymbol{x}\|_p \qquad \forall\, \boldsymbol{x}\in\mathbb{R}^m\,.$$

Consider the case $p=1, p=2$ and $p=\infty$ respectively.

1. $p=2$: Let $(\cdot,\cdot)_{\mathbb{R}^k}$ denote the inner product in Euclidean space $\mathbb{R}^k$. Then

$$\|A\boldsymbol{x}\|_2^2 = (A\boldsymbol{x}, A\boldsymbol{x})_{\mathbb{R}^n} = (\boldsymbol{x}, A^{\mathrm{T}}A\boldsymbol{x})_{\mathbb{R}^m} = (\boldsymbol{x}, P\Lambda P^{\mathrm{T}}\boldsymbol{x})_{\mathbb{R}^m} = (P^{\mathrm{T}}\boldsymbol{x}, \Lambda P^{\mathrm{T}}\boldsymbol{x})_{\mathbb{R}^n}\,,$$

in which we use the fact that $A^{\mathrm{T}}A$ is symmetric; thus diagonalizable by an orthonormal matrix P (that is, $A^{\mathrm{T}}A = P\Lambda P^{\mathrm{T}}$, $P^{\mathrm{T}}P = I$, Λ is a diagonal matrix with non-negative entries since $A^{\mathrm{T}}A$ is positive semi-definite). Let $\boldsymbol{y} = P^{\mathrm{T}}\boldsymbol{x}$. Since P is orthonormal, $\|\boldsymbol{x}\|_2 = 1$ if and only if $\|\boldsymbol{y}\|_2 = 1$; thus

$$\begin{aligned}\sup_{\|\boldsymbol{x}\|_2=1} \|A\boldsymbol{x}\|_2^2 &= \sup_{\|\boldsymbol{x}\|_2=1} (P^{\mathrm{T}}\boldsymbol{x}, \Lambda P^{\mathrm{T}}\boldsymbol{x}) = \sup_{\|\boldsymbol{y}\|_2=1} (\boldsymbol{y}, \Lambda\boldsymbol{y}) \\ &= \sup_{\|\boldsymbol{y}\|_2=1} (\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_m y_m^2) \\ &= \max\big\{\lambda_1, \cdots, \lambda_m\big\} = \text{maximum eigenvalue of } A^{\mathrm{T}}A\end{aligned}$$

which implies that $\|A\|_2 = \sqrt{\text{maximum eigenvalue of } A^{\mathrm{T}}A}$.

2. $p=\infty$: In this case we will show that

$$\|A\|_\infty = \sup_{\|\boldsymbol{x}\|_\infty=1} \|A\boldsymbol{x}\|_\infty = \max\left\{\sum_{j=1}^m |a_{1j}|, \sum_{j=1}^m |a_{2j}|, \cdots \sum_{j=1}^m |a_{nj}|\right\} = \max_{1\leqslant i\leqslant n} \sum_{j=1}^m |a_{ij}|.$$

Reason: Let $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ and $A=\big[a_{ij}\big]_{n\times m}$ (W.L.O.G. we can assume that A is not zero matrix). Then

$$A\boldsymbol{x} = \begin{bmatrix} a_{11}x_1+\cdots+a_{1m}x_m \\ a_{21}x_1+\cdots+a_{2m}x_m \\ \vdots \\ a_{n1}x_1+\cdots+a_{nm}x_m \end{bmatrix}.$$

If $\|\boldsymbol{x}\|_\infty=1$, then for each $1\leqslant i\leqslant n$,

$$|a_{i1}x_1+a_{i2}x_2+\cdots a_{im}x_m| \leqslant \sum_{j=1}^m |a_{ij}| \leqslant \max_{1\leqslant i\leqslant n}\sum_{j=1}^m |a_{ij}|;$$

thus the absolute value of each component of $A\boldsymbol{x}$, under the constraint $\|\boldsymbol{x}\|_\infty=1$, has an upper bound $\displaystyle\max_{1\leqslant i\leqslant n}\sum_{j=1}^m |a_{ij}|$. Therefore,

$$\|A\|_\infty = \sup_{\|\boldsymbol{x}\|_\infty=1}\|A\boldsymbol{x}\|_\infty = \sup_{\|\boldsymbol{x}\|_\infty=1}\max_{1\leqslant i\leqslant n}|a_{i1}x_1+a_{i2}x_2+\cdots a_{im}x_m| \leqslant \max_{1\leqslant i\leqslant n}\sum_{j=1}^m |a_{ij}|. \tag{2.2.1}$$

On the other hand, assume $\displaystyle\max_{1\leqslant i\leqslant n}\sum_{j=1}^m |a_{ij}| = \sum_{j=1}^m |a_{kj}|$ for some $1\leqslant k\leqslant n$. Let

$$\boldsymbol{x} = \big(\mathrm{sgn}(a_{k1}),\mathrm{sgn}(a_{k2}),\cdots,\mathrm{sgn}(a_{kn})\big).$$

Then $\|\boldsymbol{x}\|_\infty=1$ (since A is not zero matrix), and $\|A\boldsymbol{x}\|_\infty = \sum\limits_{j=1}^m |a_{kj}|$; thus

$$\|A\|_\infty = \sup_{\|\boldsymbol{x}\|_\infty=1}\|A\boldsymbol{x}\|_\infty \geqslant \sum_{j=1}^m |a_{kj}| = \max_{1\leqslant i\leqslant n}\sum_{j=1}^m |a_{ij}|. \tag{2.2.2}$$

The combination of (2.2.1) and (2.2.2) implies that

$$\|A\|_\infty = \max\left\{\sum_{j=1}^m |a_{1j}|, \sum_{j=1}^m |a_{2j}|, \cdots \sum_{j=1}^m |a_{nj}|\right\}. \tag{2.2.3}$$

In other words, $\|A\|_\infty$ is the largest sum of the absolute value of row entries.

3. $p = 1$: $\|A\|_1 = \max\left\{\sum_{i=1}^{n}|a_{i1}|, \sum_{i=1}^{n}|a_{i2}|, \cdots, \sum_{i=1}^{n}|a_{im}|\right\}$. This result is left as an exercise.

In general, we can also define

$$\|A\|_{p,q} = \sup_{\|\boldsymbol{x}\|_p=1}\|A\boldsymbol{x}\|_q = \sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\|A\boldsymbol{x}\|_q}{\|\boldsymbol{x}\|_p}.$$

Then $\|A\|_p = \|A\|_{p,p}$.

Example 2.20. For $1 \leqslant p < \infty$, let ℓ^p denote the collection of all sequences $\{x_n\}_{n=1}^{\infty}$ in $\mathbb{R}$ satisfying $\sum\limits_{n=1}^{\infty}|x_n|^p < \infty$; that is

$$\ell^p \equiv \left\{\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{R} \,\middle|\, \text{the series } \sum_{n=1}^{\infty}|x_n|^p \text{ converges}\right\}.$$

Then ℓ^p is a vector space over $\mathbb{R}$. The function $\|\cdot\| : \ell^p \to \mathbb{R}$ defined by $\big\|\{x_n\}_{n=1}^{\infty}\big\| = \Big(\sum\limits_{n=1}^{\infty}|x_n|^p\Big)^{\frac{1}{p}}$ is a norm on ℓ^p.

Example 2.21. Let $\mathscr{C}([a,b];\mathbb{R})$ be the collection of all continuous real-valued functions on the interval $[a,b]$; that is,

$$\mathscr{C}([a,b];\mathbb{R}) = \big\{f : [a,b] \to \mathbb{R} \,\big|\, f \text{ is continuous on } [a,b]\big\}.$$

For each $f \in \mathscr{C}([a,b];\mathbb{R})$, we define

$$\|f\|_p = \begin{cases} \Big[\displaystyle\int_a^b |f(x)|^p dx\Big]^{\frac{1}{p}} & \text{if } 1 \leqslant p < \infty, \\ \displaystyle\max_{x\in[a,b]}|f(x)| & \text{if } p = \infty. \end{cases}$$

The function $\|\cdot\|_p : \mathscr{C}([a,b];\mathbb{R}) \to \mathbb{R}$ is a norm on $\mathscr{C}([a,b];\mathbb{R})$ (Minkowski's inequality).

Definition 2.22. An ***inner product space*** $\big(\mathcal{V}, \langle\cdot,\cdot\rangle\big)$ is a vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, associated with a function $\langle\cdot,\cdot\rangle : \mathcal{V}\times\mathcal{V} \to \mathbb{F}$ such that

(1) $\langle\boldsymbol{x},\boldsymbol{x}\rangle \geqslant 0$, $\forall\, \boldsymbol{x} \in \mathcal{V}$.

(2) $\langle\boldsymbol{x},\boldsymbol{x}\rangle = 0$ if and only if $\boldsymbol{x} = 0$.

(3) $\langle\boldsymbol{x},\boldsymbol{y}+\boldsymbol{z}\rangle = \langle\boldsymbol{x},\boldsymbol{y}\rangle + \langle\boldsymbol{x},\boldsymbol{z}\rangle$ for all $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \in \mathcal{V}$.

(4) $\langle \lambda \boldsymbol{x}, \boldsymbol{y} \rangle = \lambda \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ for all $\lambda \in \mathbb{F}$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$.

(5) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$, where $\bar{c}$ denotes the complex conjugate of c.

A function $\langle \cdot, \cdot \rangle$ satisfying (1)-(5) is called an ***inner product*** on $\mathcal{V}$.

Example 2.23. Let $(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be defined by

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} x_i y_i \quad \forall\, \boldsymbol{x} = (x_1, \cdots, x_n), \boldsymbol{y} = (y_1, \cdots, y_n).$$

Then $(\cdot, \cdot)$ is an inner product on $\mathbb{R}^n$. Moreover, $\langle \cdot, \cdot \rangle : \mathbb{C}^n \to \mathbb{C}$ defined by

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i \overline{y_i} \qquad \forall\, \boldsymbol{x} = (x_1, \cdots, x_n), \boldsymbol{y} = (y_1, \cdots, y_n)$$

is an inner product on $\mathbb{C}^n$.

Example 2.24. Let $\mathscr{C}([a, b]; \mathbb{R})$ be defined as in Example 2.21. Define

$$\langle f, g \rangle = \int_a^b f(x) g(x)\, dx\,.$$

Then $\langle \cdot, \cdot \rangle : \mathscr{C}([a, b]; \mathbb{R}) \times \mathscr{C}([a, b]; \mathbb{R}) \to \mathbb{R}$ satisfies all the properties that an inner product has. Note that $\langle f, f \rangle = \|f\|_2^2$.

Similar to the inner product given above, one can also consider an inner product on $\mathscr{C}([a, b]; \mathbb{C})$, where $\mathscr{C}([a, b]; \mathbb{C})$ denotes the collection of continuous complex-valued functions defined on $[a, b]$. Note that $\mathscr{C}([a, b]; \mathbb{C})$ is a vector space over $\mathbb{R}$ and over $\mathbb{C}$, and we always viewed $\mathscr{C}([a, b], \mathbb{C})$ as a vector space over $\mathbb{C}$. On $\mathscr{C}([a, b]; \mathbb{C})$, define

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)}\, dx\,.$$

Then $\langle \cdot, \cdot \rangle : \mathscr{C}([a, b]; \mathbb{C}) \times \mathscr{C}([a, b]; \mathbb{C}) \to \mathbb{C}$ satisfies all the properties that an inner product has.

Proposition 2.25. *Let $\langle \cdot, \cdot \rangle$ be an inner product on a vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$.*

1. $\langle \lambda \boldsymbol{v} + \mu \boldsymbol{w}, \boldsymbol{u} \rangle = \lambda \langle \boldsymbol{v}, \boldsymbol{u} \rangle + \mu \langle \boldsymbol{w}, \boldsymbol{u} \rangle$ *for all* $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$ *and* $\lambda, \mu \in \mathbb{F}$.

2. $\langle \boldsymbol{u}, \lambda \boldsymbol{v} + \mu \boldsymbol{w} \rangle = \bar{\lambda} \langle \boldsymbol{u}, \boldsymbol{v} \rangle + \bar{\mu} \langle \boldsymbol{u}, \boldsymbol{w} \rangle$ *for all* $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$ *and* $\lambda, \mu \in \mathbb{F}$.

3. $\langle \mathbf{0}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \mathbf{0} \rangle = 0$ *for all* $\boldsymbol{w} \in \mathcal{V}$.

Theorem 2.26. *The inner product $\langle \cdot, \cdot \rangle$ on a vector space $\mathcal{V}$ induces a norm $\|\cdot\|$ given by $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$ and satisfies the* ***Cauchy-Schwarz inequality***

$$\big|\langle \boldsymbol{x}, \boldsymbol{y} \rangle\big| \leqslant \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\| \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}. \tag{2.2.4}$$

Moreover, for non-zero vectors $\boldsymbol{x}, \boldsymbol{y}$, the equality holds if and only if there exists $\gamma \in \mathbb{F}$ such that $\boldsymbol{x} = \gamma \boldsymbol{y}$.

Proof. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$. Define $\alpha = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$. W.L.O.G. we can assume that $\alpha \neq 0$ (for otherwise (2.2.4) holds trivially). Then there exists $\beta \in \mathbb{F}$ such that $\alpha \cdot \beta = |\alpha|$ (so $|\beta| = 1$). For any $\lambda \in \mathbb{R}$,

$$\begin{aligned} 0 &\leqslant \langle \lambda\beta\boldsymbol{x} + \boldsymbol{y}, \lambda\beta\boldsymbol{x} + \boldsymbol{y} \rangle = \lambda^2|\beta|^2\|\boldsymbol{x}\|^2 + \langle \lambda\beta\boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{y}, \lambda\beta\boldsymbol{x} \rangle + \|\boldsymbol{y}\|^2 \\ &= \lambda^2\|\boldsymbol{x}\|^2 + \lambda\beta\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \lambda\overline{\langle \beta\boldsymbol{x}, \boldsymbol{y} \rangle} + \|\boldsymbol{y}\|^2 \\ &= \lambda^2\|\boldsymbol{x}\|^2 + 2\lambda\big|\langle \boldsymbol{x}, \boldsymbol{y} \rangle\big| + \|\boldsymbol{y}\|^2. \end{aligned} \tag{2.2.5}$$

Since the right-hand side in the inequality above is always non-negative for all real λ, we must have

$$\big|\langle \boldsymbol{x}, \boldsymbol{y} \rangle\big|^2 - \|\boldsymbol{x}\|^2 \cdot \|\boldsymbol{y}\|^2 \leqslant 0$$

which implies (2.2.4).

It should be clear that (a)-(c) in Definition 2.15 are satisfied. To show that $\|\cdot\|$ satisfies the triangle inequality, by (2.2.4) we find that

$$\begin{aligned} \big(\|\boldsymbol{x}\| + \|\boldsymbol{y}\|\big)^2 - \|\boldsymbol{x} + \boldsymbol{y}\|^2 &= \|\boldsymbol{x}\|^2 + 2\|\boldsymbol{x}\|\|\boldsymbol{y}\| + \|\boldsymbol{y}\|^2 - \langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y} \rangle \\ &= 2\big(\|\boldsymbol{x}\|\|\boldsymbol{y}\| - \mathrm{Re}\langle \boldsymbol{x}, \boldsymbol{y} \rangle\big) \geqslant 2\big(\|\boldsymbol{x}\|\|\boldsymbol{y}\| - \big|\langle \boldsymbol{x}, \boldsymbol{y} \rangle\big|\big) \geqslant 0; \end{aligned}$$

thus the triangle inequality is also valid.

Finally, suppose that $\boldsymbol{x}, \boldsymbol{y} \neq \boldsymbol{0}$ and $\big|\langle \boldsymbol{x}, \boldsymbol{y} \rangle\big| = \|\boldsymbol{x}\|\|\boldsymbol{y}\|$. Then with $\lambda \in \mathbb{R}$ given by $\lambda = -\dfrac{\|\boldsymbol{y}\|}{\|\boldsymbol{x}\|}$, (2.2.5) shows that

$$0 \leqslant \|\lambda\beta\boldsymbol{x} + \boldsymbol{y}\|^2 = \lambda^2\|\boldsymbol{x}\|^2 + 2\lambda\|\boldsymbol{x}\|\|\boldsymbol{y}\| + \|\boldsymbol{y}\|^2 = (\lambda\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2 = 0;$$

thus $\lambda\beta\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{0}$. □

Corollary 2.27. *Let $f, g : [a, b] \to \mathbb{R}$ be continuous. Then*

$$\left|\int_a^b f(x)g(x)dx\right| \leqslant \left(\int_a^b |f(x)|^2 dx\right)^{\frac{1}{2}} \left(\int_a^b |g(x)|^2 dx\right)^{\frac{1}{2}}.$$

Example 2.28. Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ be a basis of $\mathcal{V}$; that is, every $\boldsymbol{x} \in \mathcal{V}$ can be uniquely expressed as

$$\boldsymbol{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j = x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n$$

for some n-tuple $(x_1, \cdots, x_n) \in \mathbb{F}^n$. Define $\|\boldsymbol{x}\|_2 = \Big(\sum\limits_{j=1}^{n} |x_j|^2\Big)^{\frac{1}{2}}$. Then $\|\cdot\|_2$ is a norm on $\mathcal{V}$. In fact, $\|\cdot\|_2$ is induced by the inner product

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{j=1}^{n} x_i \overline{y_i} \qquad \text{if } \boldsymbol{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j \text{ and } \boldsymbol{y} = \sum_{j=1}^{n} y_j \mathbf{e}_j.$$

It is also possible to talk about the notion of distance between points in a general set. A set with a distance function is called a metric space.

Definition 2.29. A ***metric space*** (M, d) is a set M associated with a function $d : M \times M \to \mathbb{R}$ such that

(i) $d(x, y) \geqslant 0$ for all $x, y \in M$.

(ii) $d(x, y) = 0$ if and only if $x = y$.

(iii) $d(x, y) = d(y, x)$ for all $x, y \in M$.

(iv) $d(x, y) \leqslant d(x, z) + d(z, y)$ for all $x, y, z \in M$.

A function d satisfying (i)-(iv) is called a ***metric*** on M.

Example 2.30 (Discrete metric)**.** Let M be a non-empty set. Define a function d_0 by

$$d_0(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y. \end{cases}$$

Then $d_0 : M \times M \to \mathbb{R}$ is a metric on M, and we call d_0 the discrete metric.

Example 2.31 (Bounded metric)**.** Let (M, d) be a metric space. Define a function ρ by

$$\rho(x, y) = \frac{d(x, y)}{1 + d(x, y)}.$$

Then $\rho : M \times M \to \mathbb{R}$ is also a metric on M.

Proposition 2.32. *Let (M, d) be a metric space, and A be a non-empty subset of M. Then (A, d) is a metric space.*

Proposition 2.33. *If $(\mathcal{V}, \|\cdot\|)$ is a normed vector space, then the function $d : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ defined by $d(x, y) = \|x - y\|$ is a metric on $\mathcal{V}$. In other words, $(\mathcal{V}, d)$ is a metric space, and we usually write $(\mathcal{V}, \|\cdot\|)$ as the metric space.*

Definition 2.34. Let (M, d) be a metric space. For each $x \in M$ and $r > 0$, the set

$$B(x, r) = \{y \in M \,|\, d(x, y) < r\}$$

is called the r***-ball*** about x or the ball centered at x with radius r.

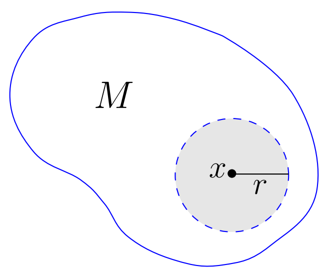


Figure 2.1: The r-ball about x in a metric space

Example 2.35. In $\mathbb{R}$, $B(x, r) = (x - r, x + r)$.

Example 2.36. Consider the 1-ball about the origin in $(\mathbb{R}^2, \|\cdot\|_p)$ for $p = 1, 2, \infty$, respectively.

1. $p = 1$: $\|\boldsymbol{x}\|_1 = |x_1| + |x_2|$, $\|\boldsymbol{x} - \boldsymbol{y}\|_1 = |x_1 - y_1| + |x_2 - y_2|$.
2. $p = 2$: $\|\boldsymbol{x}\|_2 = \sqrt{x_1^2 + x_2^2}$, $\|\boldsymbol{x} - \boldsymbol{y}\|_2 = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$.
3. $p = \infty$: $\|\boldsymbol{x}\|_\infty = \max\{|x_1|, |x_2|\}$, $\|\boldsymbol{x} - \boldsymbol{y}\|_\infty = \max\{|x_1 - y_1|, |x_2 - y_2|\}$.

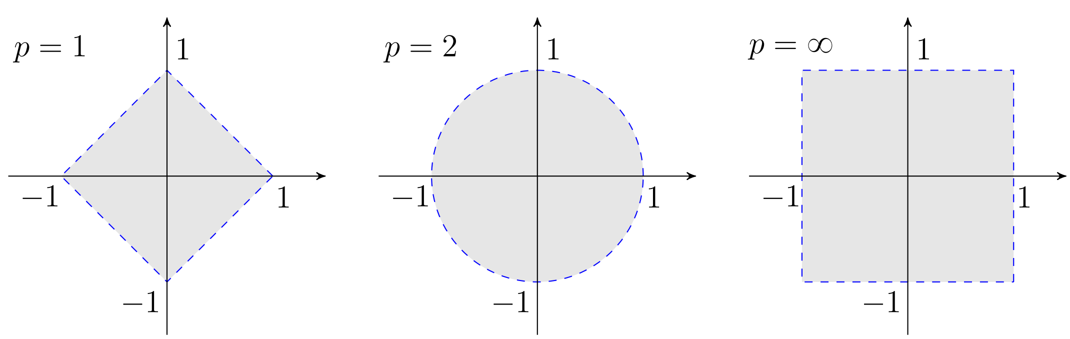


Figure 2.2: The 1-ball about $\mathbf{0}$ in $\mathbb{R}^2$ with different p

Definition 2.37. Let $\mathcal{V}$ be a vector space. Two norms $\|\cdot\|_1$, $\|\cdot\|_2$ on $\mathcal{V}$ are said to be ***equivalent*** if there exist two positive constant C, c such that

$$c\|\boldsymbol{x}\|_1 \leqslant \|\boldsymbol{x}\|_2 \leqslant C\|\boldsymbol{x}\|_1 \qquad \forall\, x \in \mathcal{V}.$$

Note that the constant c and C must be independent of $\boldsymbol{x}$.

Example 2.38. For $1 \leqslant p, q \leqslant \infty$, the p-norm $\|\cdot\|_p$ and q-norm $\|\cdot\|_q$ on $\mathbb{R}^n$ are equivalent; however, the p-norm $\|\cdot\|_p$ and the q-norm $\|\cdot\|_p$ on $\mathscr{C}([a,b];\mathbb{R})$ are **NOT** equivalent. The result is left as an exercise.

Theorem 2.39. *Let $\mathcal{V}$ be a vector space (over field $\mathbb{F}$), and $\|\cdot\|_1$, $\|\cdot\|_2$ are equivalent norms on $\mathcal{V}$. Then every ball in $(\mathcal{V}, \|\cdot\|_1)$ contains some balls in $(\mathcal{V}, \|\cdot\|_2)$ and is contained in some balls in $(\mathcal{V}, \|\cdot\|_2)$.*

Proof. Since $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, there exist positive constants c and C such that

$$c\|\boldsymbol{x}\|_1 \leqslant \|\boldsymbol{x}\|_2 \leqslant C\|\boldsymbol{x}\|_1 \qquad \forall\, x \in \mathcal{V}.$$

Let $B_1(\boldsymbol{x}, r) = \big\{\boldsymbol{y} \in \mathcal{V} \,\big|\, \|\boldsymbol{y}-\boldsymbol{x}\|_1 < r\big\}$ be a ball in $(\mathcal{V}, \|\cdot\|_1)$. Let $\delta = cr$ and $R = Cr$. Then with $B_2(\boldsymbol{x}, r)$ denoting the set $\big\{\boldsymbol{y} \in \mathcal{V} \,\big|\, \|\boldsymbol{y}-\boldsymbol{x}\|_2 < r\big\}$, we have

$$\|\boldsymbol{y}-\boldsymbol{x}\|_1 \leqslant \frac{1}{c}\|\boldsymbol{y}-\boldsymbol{x}\|_2 < r \quad \forall\, \boldsymbol{y} \in B_2(\boldsymbol{x}, \delta) \quad \text{and} \quad \|\boldsymbol{y}-\boldsymbol{x}\|_2 \leqslant C\|\boldsymbol{y}-\boldsymbol{x}\|_1 < R$$

In other words, $B_2(\boldsymbol{x}, \delta) \subseteq B_1(\boldsymbol{x}, r)$ and $B_1(\boldsymbol{x}, r) \subseteq B_2(\boldsymbol{x}, R)$. □

2.3 Sequences in Metric Spaces

2.3.1 Sequences

Recall that a ***sequence*** in a set S is a function $f : \mathbb{N} \to S$, and $f(n)$ is called the n-th terms of the sequence. A sequence in S is usually denoted by $\big\{f(n)\big\}_{n=1}^{\infty}$ or $\{x_n\}_{n=1}^{\infty}$ with $x_n = f(n)$.

Definition 2.40. Let (M, d) be a metric space. A sequence $\{x_n\}_{n=1}^{\infty} \subseteq M$ is said to be ***convergent*** if there exists $x \in M$ such that for every $\varepsilon > 0$, there exists $N > 0$ such that

$$d(x_n, x) < \varepsilon \quad \text{whenever} \quad n \geqslant N.$$

Such an x is called a ***limit*** of the sequence. In notation,

$$\{x_n\}_{n=1}^{\infty} \subseteq M \text{ is convergent } \Leftrightarrow (\exists\, x \in M)(\forall\, \varepsilon > 0)(\exists\, N > 0)\big(n \geqslant N \Rightarrow d(x_n, x) < \varepsilon\big).$$

If x is a limit of $\{x_n\}_{n=1}^{\infty}$, we say $\{x_n\}_{n=1}^{\infty}$ converges to x and write $x_n \to x$ as $n \to \infty$. If no such x exists we say that $\{x_n\}_{n=1}^{\infty}$ ***diverges*** or $\lim\limits_{n\to\infty} x_n$ does not exist.

Remark 2.41. Similar to Definition 1.26, the convergence of a sequence $\{x_n\}_{n=1}^{\infty}$ in a metric space can be stated as follows: a sequence $\{x_n\}_{n=1}^{\infty} \subseteq M$ is said to be ***convergent*** if there exists $x \in M$ such that for every $\varepsilon > 0$, there exists $N > 0$ such that

$$\#\big\{n \in \mathbb{N} \,\big|\, x_n \notin B(x, \varepsilon)\big\} < \infty.$$

Similar to Proposition 1.28, we have the following

Proposition 2.42. *Let (M, d) be a metric space. If $\{x_n\}_{n=1}^{\infty}$ is a sequence in M, and $x_n \to x$ and $x_n \to y$ as $n \to \infty$, then $x = y$. (The uniqueness of the limit).*

Proof. Assume the contrary that $x \neq y$. Then $\varepsilon = \dfrac{d(x, y)}{2} > 0$. Then there exist $N_1, N_2 > 0$ such that $d(x_n, x) < \varepsilon$ for all $n \geqslant N_1$ and $d(x_n, y) < \varepsilon$ for all $n \geqslant N_2$. Let $N = \max\{N_1, N_2\}$. Then if $n \geqslant N$,

$$d(x, y) \leqslant d(x, x_n) + d(x_n, y) < 2\varepsilon = d(x, y),$$

a contradiction. □

Notation: Similar to the notation used to denote the unique limit of a convergent sequence in an ordered field, we also use $\lim\limits_{n\to\infty} x_n$ to denote the limit of a convergent sequence $\{x_n\}_{n=1}^{\infty} \subseteq M$.

Remark 2.43. Similar to Remark 1.31, the proposition above implies that $x_n \to x$ as $n \to \infty$ if and only if $\lim\limits_{n\to\infty} d(x_n, x) = 0$.

Remark 2.44. Let (M, d) be a metric space. A sequence $\{x_n\}_{n=1}^{\infty} \subseteq M$ diverges if (and only if)

$$\forall\, x \in M, \exists\, \varepsilon > 0 \ni \forall\, N > 0, \exists\, n \geqslant N \text{ such that } d(x_n, x) \geqslant \varepsilon.$$

Definition 2.45. Let (M, d) be a metric space. A sequence $\{x_n\}_{n=1}^{\infty} \subseteq M$ is said to be ***bounded***（有界的）if there exist $y \in M$ and $r > 0$ such that $d(x_n, y) < r$ for all $n \in \mathbb{N}$. In other words, sequence $\{x_n\}_{n=1}^{\infty}$ is bounded if it is contained in some r-ball.

Remark 2.46. In a normed vector space $(\mathcal{V}, \|\cdot\|)$, the boundedness of a sequence $\{\boldsymbol{x}_n\}_{n=1}^\infty$ is equivalent to that there exists $r > 0$ such that $\|\boldsymbol{x}_n\| < r$ for all $n \in \mathbb{N}$. In other words, the point $\boldsymbol{y}$ in the definition above is the zero vector.

Proposition 2.47. *A convergent sequence is bounded*（收斂數列必有界）.

Proof. Let $\{x_n\}_{n=1}^\infty$ be a convergent sequence with limit x. Then there exists $N > 0$ such that $x_n \in B(x, 1)$ whenever $n \geqslant N$, or equivalently,

$$d(x_n, x) < 1 \quad \text{whenever} \quad n \geqslant N.$$

Let $r = \max\big\{d(x_1, x), d(x_2, x), \cdots, d(x_{N-1}, x)\big\} + 1$. Then $d(x_n, x) < r$ for all $n \in \mathbb{N}$. □

Theorem 2.48. *Let $(\mathcal{V}, \|\cdot\|)$ be a normed vector space over a scalar field $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$), $\{\boldsymbol{x}_n\}_{n=1}^\infty$, $\{\boldsymbol{y}_n\}_{n=1}^\infty$ be sequences in $\mathcal{V}$, and $\{\lambda_n\}_{n=1}^\infty$ be a sequence in $\mathbb{F}$. Suppose that $\boldsymbol{x}_n \to \boldsymbol{x}$, $\boldsymbol{y}_n \to \boldsymbol{y}$ and $\lambda_n \to \lambda$ as $n \to \infty$. Then*

1. $\boldsymbol{x}_n \pm \boldsymbol{y}_n \to \boldsymbol{x} \pm \boldsymbol{y}$ *as* $n \to \infty$.
2. $\lambda_n \boldsymbol{x}_n \to \lambda \boldsymbol{x}$ *as* $n \to \infty$.
3. *If* $\lambda_n, \lambda \neq 0$, *then* $\dfrac{x_n}{\lambda_n} \to \dfrac{x}{\lambda}$ *as* $n \to \infty$.

If in addition that $\mathcal{V}$ is an inner product space equipped with inner product $\langle \cdot, \cdot \rangle$ which induces the norm $\|\cdot\|$, then

4. $\langle \boldsymbol{x}_n, \boldsymbol{y}_n \rangle \to \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ *as* $n \to \infty$.

Proof. We only prove 3 and 4. The proof of 1 and 2 are left as an exercise.

3. It suffices to show that $\lim\limits_{n\to\infty} \dfrac{1}{\lambda_n} = \dfrac{1}{\lambda}$ if $\lambda_n, \lambda \neq 0$ (because of 2). Since $\lim\limits_{n\to\infty} \lambda_n = \lambda$, there exists $N_1 > 0$ such that $|\lambda_n - \lambda| < \dfrac{|\lambda|}{2}$ whenever $n \geqslant N_1$. Therefore, $|\lambda| - |\lambda_n| < \dfrac{|\lambda|}{2}$ for all $n \geqslant N_1$ which further implies that $|\lambda_n| > \dfrac{|\lambda|}{2}$ for all $n \geqslant N_1$.

   Let $\varepsilon > 0$ be given. Since $\lim\limits_{n\to\infty} \lambda_n = \lambda$, there exists $N_2 > 0$ such that $|\lambda_n - \lambda| < \dfrac{|\lambda|^2}{2}\varepsilon$ whenever $n \geqslant N_2$. Define $N = \max\{N_1, N_2\}$. Then if $n \geqslant N$,

$$\Big|\frac{1}{\lambda_n} - \frac{1}{\lambda}\Big| = \frac{|\lambda_n - \lambda|}{|\lambda_n||\lambda|} < \frac{|\lambda|^2}{2}\varepsilon \cdot \frac{1}{|\lambda|}\frac{2}{|\lambda|} = \varepsilon.$$

4. Let $\varepsilon > 0$ be given. Since $\boldsymbol{x}_n \to \boldsymbol{x}$ and $\boldsymbol{y}_n \to \boldsymbol{y}$ as $n \to \infty$, by Proposition 2.47 and Remark 2.46 there exists $M > 0$ such that $\|\boldsymbol{x}_n\| \leqslant M$ and $\|\boldsymbol{y}_n\| \leqslant M$. Moreover,

$$\exists\, N_1 > 0 \ni \|\boldsymbol{x}_n - \boldsymbol{x}\| < \frac{\varepsilon}{2M} \quad \text{whenever} \quad n \geqslant N_1$$

and

$$\exists\, N_2 > 0 \ni \|\boldsymbol{y}_n - \boldsymbol{y}\| < \frac{\varepsilon}{2M} \quad \text{whenever} \quad n \geqslant N_2\,.$$

Define $N = \max\{N_1, N_2\}$. Then if $n \geqslant N$,

$$\begin{aligned} \big|\langle \boldsymbol{x}_n, \boldsymbol{y}_n\rangle - \langle \boldsymbol{x}, \boldsymbol{y}\rangle\big| &= \big|\langle \boldsymbol{x}_n, \boldsymbol{y}_n\rangle - \langle \boldsymbol{x}_n, \boldsymbol{y}\rangle + \langle \boldsymbol{x}_n, \boldsymbol{y}\rangle - \langle \boldsymbol{x}, \boldsymbol{y}\rangle\big| \\ &\leqslant \big|\langle \boldsymbol{x}_n, \boldsymbol{y}_n - \boldsymbol{y}\rangle\big| + \big|\langle \boldsymbol{y}, \boldsymbol{x}_n - \boldsymbol{x}\rangle\big| \\ &\leqslant M\|\boldsymbol{y}_n - \boldsymbol{y}\| + M\|\boldsymbol{x}_n - \boldsymbol{x}\| < M \cdot \frac{\varepsilon}{2M} + M \cdot \frac{\varepsilon}{2M} = \varepsilon\,. \end{aligned}$$ □

Proposition 2.49. *In $\mathbb{R}^n$, a sequence of vectors converges if and only if every component of the vectors converges. In other words, in $\mathbb{R}^n$*

Componentwise convergence $\Leftrightarrow$ *Convergence.*

Proof. Let $\{\boldsymbol{v}_k\}_{k=1}^{\infty}$, $\boldsymbol{v}_k = (v_k^{(1)}, v_k^{(2)}, \cdots, v_k^{(n)})$, be a sequence of vectors in $\mathbb{R}^n$.

"$\Leftarrow$" Suppose that $\boldsymbol{v}_k \to \boldsymbol{v} = (v^{(1)}, \cdots, v^{(n)})$ as $k \to \infty$. Let $\varepsilon > 0$ be given. There exists $N > 0$ such that

$$\|\boldsymbol{v}_k - \boldsymbol{v}\|_2 < \varepsilon \quad \text{whenever} \quad k \geqslant N\,;$$

thus if $k \geqslant N$,

$$|v_k^{(i)} - v^{(i)}| \leqslant \sqrt{(v_k^{(1)} - v^{(1)})^2 + \cdots + (v_k^{(n)} - v^{(n)})^2} = \|\boldsymbol{v}_k - \boldsymbol{v}\|_2 < \varepsilon.$$

"$\Rightarrow$" Suppose that $v_k^{(i)} \to v^{(i)}$ as $k \to \infty$ for each $1 \leqslant i \leqslant n$. Let $\varepsilon > 0$ be given. For each $1 \leqslant i \leqslant n$, there exist $N_i > 0$ such that

$$\big|v_k^{(i)} - v^{(i)}\big| < \frac{\varepsilon}{\sqrt{n}} \quad \text{whenever} \quad k \geqslant N_i\,.$$

Let $\boldsymbol{v} = \big(v^{(1)}, v^{(2)}, \cdots, v^{(n)}\big)$ and $N = \max\{N_1, N_2, \cdots, N_n\}$. Then if $k \geqslant N$,

$$\|\boldsymbol{v}_k - \boldsymbol{v}\|_2 = \sqrt{(v_k^{(1)} - v^{(1)})^2 + \cdots + (v_k^{(n)} - v^{(n)})^2} < \sqrt{\frac{\varepsilon^2}{n} + \cdots + \frac{\varepsilon^2}{n}} = \varepsilon.$$ □

Example 2.50. Let $\boldsymbol{v}_k = \big(\frac{1}{k}, \frac{1}{k^2}\big) \in \mathbb{R}^2$. Then $\boldsymbol{v}_k \to (0,0)$ as $k \to \infty$ since

$$\sqrt{(\frac{1}{k} - 0)^2 + (\frac{1}{k^2} - 0)^2} = \frac{1}{k^2}\sqrt{k^2+1} \to 0 \text{ as } k \to \infty\,.$$

On the other hand, since $\frac{1}{k} \to 0$ and $\frac{1}{k^2} \to 0$ as $k \to \infty$, Proposition 2.49 implies that $\boldsymbol{v}_k \to (0,0)$ as $k \to \infty$.

Theorem 2.51 (Bolzano-Weierstrass)**.** *Every bounded sequence in* $(\mathbb{R}^n, \|\cdot\|_2)$ *has a convergent subsequence.*

Proof. We prove by induction. Let

$$S = \Big\{n \in \mathbb{N} \,\Big|\, \text{every bounded sequence in } (\mathbb{R}^n, \|\cdot\|_2) \text{ has a convergent subsequence}\Big\}.$$

Then $1 \in S$ because of the Bolzano-Weierstrass Property of $\mathbb{R}$.

Suppose that $n \in S$. Let $\{\boldsymbol{x}_k\}_{k=1}^\infty$ be a bounded sequence in $\mathbb{R}^{n+1}$. Write $\boldsymbol{x}_k = \big(x_k^{(1)}, x_k^{(2)}, \cdots, x_k^{(n)}, x_k^{(n+1)}\big)$, and let $\boldsymbol{y}_k = \big(x_k^{(1)}, x_k^{(2)}, \cdots, x_k^{(n)}\big)$. Since $\{\boldsymbol{x}_k\}_{k=1}^\infty$ is bounded, there exists $M > 0$ such that $\|\boldsymbol{x}_k\|_2 \leqslant M$ for all $k \in \mathbb{N}$; thus

$$\|\boldsymbol{y}_k\|_2 \leqslant M \quad \forall\, k \in \mathbb{N} \qquad \text{and} \qquad \big|x_k^{(n+1)}\big| \leqslant M \qquad \forall\, k \in \mathbb{N}.$$

that is, $\{\boldsymbol{y}_k\}_{k=1}^\infty$ is bounded in $\mathbb{R}^n$ and $\big\{x_k^{(n+1)}\big\}_{k=1}^n$ is bounded in $\mathbb{R}$. By the assumption that $n \in S$, $\{y_k\}_{k=1}^\infty$ has a convergent subsequence $\{\boldsymbol{y}_{k_j}\}_{j=1}^\infty$ of $\{\boldsymbol{y}_k\}_{k=1}^\infty$ which converges to $\boldsymbol{y} = \big(y^{(1)}, y^{(2)}, \cdots, y^{(n)}\big)$. Applying the Bolzano-Weierstrass Property to the sequence $\big\{x_{k_j}^{(n+1)}\big\}_{j=1}^\infty$, we obtain a subsequence $\big\{x_{k_{j_\ell}}^{(n+1)}\big\}_{\ell=1}^\infty$ of $\big\{x_{k_j}^{(n+1)}\big\}_{k=1}^\infty$ so that $x_{k_{j_\ell}}^{(n+1)} \to y^{(n+1)}$ as $\ell \to \infty$. Let $\boldsymbol{z}_\ell = \boldsymbol{x}_{k_{j_\ell}}$. Then $\{\boldsymbol{z}_\ell\}_{\ell=1}^\infty$ is a subsequence of $\{\boldsymbol{x}_k\}_{k=1}^\infty$ and $\{\boldsymbol{z}_\ell\}_{\ell=1}^\infty$ converges to $\boldsymbol{z} = \big(y^{(1)}, y^{(2)}, \cdots, y^{(n+1)}\big)$ by Proposition 2.49. Therefore, $n+1 \in S$.

By induction, $S = \mathbb{N}$; thus the theorem is proved. □

Remark 2.52. By identifying $\mathbb{C}^n$ as $\mathbb{R}^{2n}$, Theorem 2.51 also implies that every bounded sequence in $\mathbb{C}^n$ has a convergent subsequence.

Definition 2.53. Let (M, d) be a metric space, and $\{x_n\}_{n=1}^\infty$ be a sequence in M. A point $x \in M$ is called a ***cluster point*** of $\{x_n\}_{n=1}^\infty$ if

$$\forall\, \varepsilon > 0,\ \#\big\{n \in \mathbb{N} \,\big|\, x_n \in B(x, \varepsilon)\big\} = \infty\,.$$

Example 2.54. Let $x_n = (-1)^n$. Then 1 and -1 are the only two cluster points of $\{x_n\}_{n=1}^\infty$.

Example 2.55. Let $x_n = (-1)^n + \dfrac{1}{n}$. Then 1 and -1 are cluster points of $\{x_n\}_{n=1}^\infty$: Let $\varepsilon > 0$ be given. We observe that

$$\big\{n \in \mathbb{N} \,\big|\, x_n \in (1-\varepsilon, 1+\varepsilon)\big\} \supseteq \Big\{n \in \mathbb{N} \,\Big|\, n \text{ is even}, \frac{1}{n} < \varepsilon\Big\};$$

thus $\#\big\{n \in \mathbb{N} \,\big|\, x_n \in (1-\varepsilon, 1+\varepsilon)\big\} = \infty$. Similarly, -1 is a cluster point.

On the other hand, each $a \neq \pm 1$ is not a cluster point of $\{x_n\}_{n=1}^\infty$.

Let $\{x_n\}_{n=1}^\infty$ be a sequence. Recall that a subsequence of $\{x_n\}_{n=1}^\infty$ is a sequence $\{y_j\}_{j=1}^\infty$ satisfying that $y_j = x_{f(j)}$ for some strictly increasing function $f : \mathbb{N} \to \mathbb{N}$. In other words, each strictly increasing function $f : \mathbb{N} \to \mathbb{N}$ corresponds to a subsequence of $\{x_n\}_{n=1}^\infty$ and vice versa.

Similar to Proposition 1.68, we have the following

Proposition 2.56. *Let (M, d) be a metric space, $\{x_n\}_{n=1}^\infty$ be a sequence in M, and $x \in M$.*

1. *x is a cluster point of $\{x_n\}_{n=1}^\infty$ if and only if for each $\varepsilon > 0$ and $N > 0$, there exists $n \geqslant N$ such that $d(x_n, x) < \varepsilon$.*

2. *x is a cluster point of $\{x_n\}_{n=1}^\infty$ if and only if there exists a subsequence $\{x_{n_j}\}_{j=1}^\infty$ of $\{x_n\}_{n=1}^\infty$ converges to x.*

3. *$x_n \to x$ as $n \to \infty$ if and only if every proper subsequence of $\{x_n\}_{n=1}^\infty$ converges to x.*

4. *$x_n \to x$ as $n \to \infty$ if and only if every proper subsequence of $\{x_n\}_{n=1}^\infty$ has a further subsequence that converges to x.*

Proof. We only prove 1 and 2 since the proof of 3 and 4 are similar to the one given in Proposition 1.60.

1. "$\Rightarrow$" Let $\varepsilon > 0$ and $N > 0$ be given. Since $\#\big\{n \in \mathbb{N} \,\big|\, x_n \in B(x, \varepsilon)\big\} = \infty$, there exist natural numbers $n_1 < n_2 < n_3 < \cdots$ such that

$$\big\{n \in \mathbb{N} \,\big|\, x_n \in B(x, \varepsilon)\big\} = \big\{n_1, n_2, n_3, \cdots\big\}.$$

Note that $n_j \geqslant j$; thus $n_N \geqslant N$.

"$\Leftarrow$" Let $\varepsilon > 0$ be given. Pick $n_1 \geqslant 1$ such that $d(x_{n_1}, x) < \varepsilon$, then pick $n_2 \geqslant n_1 + 1$ such that $d(x_{n_2}, x) < \varepsilon$. We continue this process and obtain a subsequence $\{x_{n_j}\}_{j=1}^\infty$ satisfying $d(x_{n_j}, x) < \varepsilon$ for all $j \in \mathbb{N}$. Then $\big\{n \in \mathbb{N} \,\big|\, x_n \in B(x, \varepsilon)\big\} \supseteq \{n_1, n_2, \cdots\}$ which implies that $\#\big\{n \in \mathbb{N} \,\big|\, x_n \in B(x, \varepsilon)\big\} = \infty$.

2. "$\Rightarrow$" By 1, we can pick $n_1 \geqslant 1$ such that $d(x_{n_1}, x) < 1$ and then pick $n_2 \geqslant n_1 + 1$ such that $d(x_{n_2}, x) < \frac{1}{2}$. In general, we can pick $n_{k+1} \geqslant n_k + 1$ so that
$$d(x_{n_k}, x) < \frac{1}{k} \qquad \forall\, k \in \mathbb{N}.$$
Therefore, $\lim\limits_{k\to\infty} x_{n_k} = x$.
"$\Leftarrow$" Let $\varepsilon > 0$ be given. By assumption there exists $J > 0$ such that $d(x_{n_j}, x) < \varepsilon$ whenever $j \geqslant J$. Then $\big\{n \in \mathbb{N} \,\big|\, x_n \in B(x,\varepsilon)\big\} \supseteq \{n_J, n_{J+1}, \cdots\}$ which implies that $\#\big\{n \in \mathbb{N} \,\big|\, x_n \in B(x,\varepsilon)\big\} = \infty$. $\square$

2.3.2 Cauchy sequences, Banach spaces and Hilbert spaces

Definition 2.57. A sequence $\{x_n\}_{n=1}^\infty$ in a metric space (M,d) is said to be ***Cauchy*** if
$$(\forall\, \varepsilon > 0)(\exists\, N > 0)\big(n, m \geqslant N \Rightarrow d(x_n, x_m) < \varepsilon\big).$$

Similar to Proposition 1.74, Lemma 1.75 and 1.76, we have the following

Proposition 2.58. 1. *Every convergent sequence* (*in a metric space* (M,d)) *is Cauchy.*

2. *Every Cauchy sequence* (*in a metric space* (M,d)) *is bounded.*

3. *If a subsequence of Cauchy sequence* (*in a metric space* (M,d)) *converges, then this Cauchy sequence also converges.*

Proof. See the proof of Proposition 1.74 and Lemma 1.76 by changing $|x-y|$ to $d(x,y)$ with appropriate x and y. $\square$

Remark 2.59. By 2 of Proposition 2.56 and 3 of Proposition 2.58, we find that if $\{x_k\}_{k=1}^\infty$ is a Cauchy sequence but $\{x_k\}_{k=1}^\infty$ does not converge, then
$$(\forall\, y)(\exists\, r > 0)\big(\#\big\{n \in \mathbb{N} \,\big|\, x_n \in B(y,r)\big\} < \infty\big).$$

Theorem 2.60. *A sequence in* $\mathbb{R}^n$ *converges if and only if the sequence is Cauchy* (*because of the inequality* $\max\limits_{1\leqslant i\leqslant n} \big|v_k^{(i)} - v_\ell^{(i)}\big| \leqslant \|\boldsymbol{v}_k - \boldsymbol{v}_\ell\|_2 \leqslant \sqrt{n} \max\limits_{1\leqslant i\leqslant n} \big|v_k^{(i)} - v_\ell^{(i)}\big|$).

Now we would like to define the completeness of a normed vector space. Recall that in an Archimedean ordered field, the following four properties are equivalent:

1. the Bolzano-Weierstrass property (有界數列必有收斂子數列),

2. the monotone sequence property (單調有界數列必收斂),
3. the least upper bound property (非空集合有上界必有最小上界),
4. the property that every Cauchy sequence converges (柯西數列必收斂).

Since in general a normed vector space cannot be an ordered field, we cannot define the completeness through the monotone sequence property or the least upper bound property. For the completeness of normed vector spaces, we use the ***Cauchy completeness***.

Definition 2.61. A metric space (M, d) is said to be complete if every Cauchy sequence in M converges. A ***Banach space*** is a complete normed vector space, and a ***Hilbert space*** is a complete inner product space (that is, a Banach space whose norm is induced by the inner product).

Example 2.62. The Euclidean n-space $\mathbb{R}^n$, equipped with p-norm, is a Banach space for all $1 \leqslant p \leqslant \infty$. To see this, let $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$ be a Cauchy sequence in $\mathbb{R}^n$. Then the sequence $\{x_k^{(i)}\}_{k=1}^{\infty}$, consisting of the i-th components of $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$, is Cauchy for all $1 \leqslant i \leqslant n$ since

$$\left|x_k^{(i)} - x_\ell^{(i)}\right| \leqslant \|\boldsymbol{x}_k - \boldsymbol{x}_\ell\|_p \qquad \forall\, 1 \leqslant i \leqslant n \text{ and } 1 \leqslant p \leqslant \infty.$$

By the completeness of $\mathbb{R}$, the real sequence $\{x_k^{(i)}\}_{k=1}^{\infty}$ converges for all $1 \leqslant i \leqslant n$; thus each component of $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$ converges. Proposition 2.49 implies that $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$ converges.

2.4 Series of Real Numbers and Vectors

Definition 2.63. Let $(\mathcal{V}, \|\cdot\|)$ be a normed space. A series $\sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$, where $\{\boldsymbol{x}_k\}_{k=1}^{\infty} \subseteq \mathcal{V}$, is said to ***converge*** to $\boldsymbol{S} \in \mathcal{V}$ if the partial sum $\boldsymbol{S}_n = \sum\limits_{k=1}^{n} \boldsymbol{x}_k$ converges to $\boldsymbol{S}$, and one writes $\boldsymbol{S} = \sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$ if this is the case. A series in $\mathcal{V}$ is said to converge or be convergent if it converges to some element in $\mathcal{V}$.

The following proposition is a direct consequence of the monotone sequence property of the real number system $\mathbb{R}$.

Proposition 2.64. *Let $\{x_k\}_{k=1}^{\infty}$ be a sequence of real numbers, and $x_k \geqslant 0$ for all $k \in \mathbb{N}$. Then $\sum\limits_{k=1}^{\infty} x_k$ converges if and only if the sequence of partial sums $\{S_n\}_{n=1}^{\infty}$, where $S_n = \sum\limits_{k=1}^{n} x_k$, is bounded (from above).*

Theorem 2.65 (Cauchy's criterion)**.** *Let $(\mathcal{V}, \|\cdot\|)$ be a Banach space. A series $\sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$ converges if and only if*

$$\forall\, \varepsilon > 0, \exists\, N > 0 \ni \|\boldsymbol{x}_k + \boldsymbol{x}_{k+1} + \cdots + \boldsymbol{x}_{k+p}\| < \varepsilon \quad \textit{whenever} \quad k \geqslant N, p \geqslant 0\,.$$

Proof. Let $\boldsymbol{S}_n = \sum\limits_{k=1}^{n} \boldsymbol{x}_k$ be partial sum of $\sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$. Then

$$\begin{aligned}
\{\boldsymbol{S}_n\}_{n=1}^{\infty} \text{ converges in } \mathcal{V} &\Leftrightarrow \{\boldsymbol{S}_n\}_{n=1}^{\infty} \text{ is Cauchy} \\
&\Leftrightarrow \forall\, \varepsilon > 0, \exists\, N > 0 \ni \|\boldsymbol{S}_n - \boldsymbol{S}_m\| < \varepsilon \text{ whenever } n, m \geqslant N \\
&\Leftrightarrow \forall\, \varepsilon > 0, \exists\, N > 0 \ni \|\boldsymbol{x}_{n+1} + \boldsymbol{x}_{n+2} + \cdots + \boldsymbol{x}_m\| < \varepsilon \text{ whenever } m > n \geqslant N \\
&\Leftrightarrow \forall\, \varepsilon > 0, \exists\, N > 0 \ni \|\boldsymbol{x}_k + \boldsymbol{x}_{k+1} + \cdots + \boldsymbol{x}_{k+p}\| < \varepsilon \text{ whenever } k \geqslant N+1, p \geqslant 0.
\end{aligned}$$ □

Corollary 2.66 (n-th term test)**.** *If $\sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$ converges, then $\|\boldsymbol{x}_k\| \to 0$ as $k \to \infty$, and if $\|\boldsymbol{x}_k\| \not\to 0$ as $k \to \infty$, then $\sum\limits_{k=1}^{\infty} x_k$ diverges.*

Proof. Take $p = 0$ in Theorem 2.65. □

Definition 2.67. A series $\sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$ is said to ***converge absolutely*** if $\sum\limits_{k=1}^{\infty} \|\boldsymbol{x}_k\|$ converges in $\mathbb{R}$. A series that is convergent but not absolutely convergent is said to be ***conditionally convergent***.

Example 2.68. $\sum\limits_{k=1}^{\infty} \dfrac{(-1)^k}{k}$ is conditionally convergent. See Theorem 2.70 for the reason.

Theorem 2.69. *Let $(\mathcal{V}, \|\cdot\|)$ be a Banach space, and $\{\boldsymbol{x}_n\}_{n=1}^{\infty}$ be a sequence in $\mathcal{V}$. If $\sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$ converges absolutely, then $\sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$ converges.*

Proof. If $\sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$ converges absolutely, then $S_n = \sum\limits_{k=1}^{n} \|\boldsymbol{x}_k\|$ converges in $\mathbb{R}$. Then

$$\forall\, \varepsilon > 0, \exists\, N > 0 \ni \big|\|\boldsymbol{x}_k\| + \|\boldsymbol{x}_{k+1}\| + \cdots + \|\boldsymbol{x}_{k+p}\|\big| < \varepsilon \quad \text{whenever} \quad k \geqslant N, p \geqslant 0\,.$$

Therefore, if $k \geqslant N, p \geqslant 0$,

$$\|\boldsymbol{x}_k + \boldsymbol{x}_{k+1} + \cdots + \boldsymbol{x}_{k+p}\| \leqslant \|\boldsymbol{x}_k\| + \cdots + \|\boldsymbol{x}_{k+p}\| < \varepsilon\,,$$

and the convergence of $\sum\limits_{k=1}^{\infty} \boldsymbol{x}_k$ is guaranteed by the Cauchy criterion. □

Theorem 2.70. 1. ***Geometric series***:

(a) *If* $|r| < 1$, *then* $\sum\limits_{k=1}^{\infty} r^k$ *converges absolutely to* $\dfrac{r}{1-r}$.

(b) *If* $|r| \geqslant 1$, *then* $\sum\limits_{k=1}^{\infty} r^k$ *does not converge (diverge).*

2. ***Comparison test***: *Let* $\{a_k\}_{k=1}^{\infty}$ *and* $\{b_k\}_{k=1}^{\infty}$ *be sequences of real numbers.*

(a) *If* $\sum\limits_{k=1}^{\infty} a_k$ *converges, and* $0 \leqslant b_k \leqslant a_k$, *then* $\sum\limits_{k=1}^{\infty} b_k$ *converges.*

(b) *If* $\sum\limits_{k=1}^{\infty} a_k$ *diverges, and* $0 \leqslant a_k \leqslant b_k$, *then* $\sum\limits_{k=1}^{\infty} b_k$ *diverges.*

3. ***Integral test***: *If* f *is continuous, non-negative, and monotone decreasing on* $[1, \infty)$, *then* $\sum\limits_{k=1}^{\infty} f(k)$ *converges if and only if the improper integral* $\displaystyle\int_1^{\infty} f(x)dx < \infty$.

4. ***Root test***: *Let* $\{x_k\}_{k=1}^{\infty}$ *be a sequence of real numbers.*

(a) *If* $\limsup\limits_{k\to\infty} \sqrt[k]{|x_k|} < 1$, *then* $\sum\limits_{k=1}^{\infty} x_k$ *converges absolutely.*

(b) *If* $\limsup\limits_{k\to\infty} \sqrt[k]{|x_k|} > 1$, *then* $\sum\limits_{k=1}^{\infty} x_k$ *diverges.*

5. ***Ratio and comparison test***: *Let* $\{a_k\}_{k=1}^{\infty}$ *and* $\{b_k\}_{k=1}^{\infty}$ *be sequences of real numbers, and* $b_k > 0$ *for all* $k \in \mathbb{N}$.

(a) $\limsup\limits_{k\to\infty} \dfrac{|a_k|}{b_k} < \infty$, $\sum\limits_{k=1}^{\infty} b_k$ *is convergent, then* $\sum\limits_{k=1}^{\infty} a_k$ *converges absolutely.*

(b) $\liminf\limits_{k\to\infty} \dfrac{a_k}{b_k} > 0$, $\sum\limits_{k=1}^{\infty} b_k$ *is divergent, then* $\sum\limits_{k=1}^{\infty} a_k$ *diverges.*

6. ***Dirichlet test***: *Let* $\{a_k\}_{n=1}^{\infty}$, $\{p_k\}_{n=1}^{\infty}$ *be sequences of real numbers such that*

(a) *the sequence of partial sums of the series* $\sum\limits_{k=1}^{\infty} a_k$ *is bounded; that is, there exists* $M \in \mathbb{R}$ *such that* $\Big|\sum\limits_{k=1}^{n} a_k\Big| \leqslant M$ *for all* $n \in \mathbb{N}$.

(b) $\{p_k\}_{k=1}^{\infty}$ *is a decreasing sequence, and* $\lim\limits_{k\to\infty} p_k = 0$.

Then $\sum\limits_{k=1}^{\infty} a_k p_k$ *converges.*

Proof. Note that 1 follows from the fact that $\sum\limits_{k=1}^{n} r^k = \dfrac{r - r^{n+1}}{1-r}$ if $r \neq 1$, the comparison test follows from Proposition 2.64, and the integral test follows from the fact that

$$\int_1^{n+1} f(x)\,dx \leqslant S_n \equiv \sum_{k=1}^{n} f(k) \leqslant f(1) + \int_1^n f(x)\,dx$$

with an application of the comparison test, we only prove 4, 5 and 6.

4. Let $r = \limsup\limits_{k\to\infty} \sqrt[k]{|x_k|}$.

   (a) Suppose that $0 \leqslant r < 1$. By Proposition 1.98, there exists $N > 0$ such that $\sqrt[k]{|x_k|} < \dfrac{r+1}{2}\Big(= r + \dfrac{1-r}{2}\Big)$ for all $k \geqslant N$. This implies that
   $$|x_k| \leqslant \Big(\frac{r+1}{2}\Big)^k \qquad \forall\, k \geqslant N\,.$$
   Since $\Big|\dfrac{r+1}{2}\Big| < 1$, the geometric series $\sum\limits_{k=1}^{\infty} \big(\dfrac{r+1}{2}\big)^k$ converges; thus the comparison test implies that the series $\sum\limits_{k=1}^{\infty} |x_k|$ converges.

   (b) Suppose that $r > 1$. By Proposition 1.98 there exist $n_1 < n_2 < \cdots < n_j < \cdots$ such that
   $$\sqrt[k]{|x_k|} > \frac{r+1}{2} \qquad \forall\, k = n_1, n_2, \cdots,$$
   The statement above then implies that $\lim\limits_{k\to\infty} x_k$, if exists, cannot be zero; thus the n-th term test shows that $\sum\limits_{k=1}^{\infty} x_k$ diverges.

5. (a) Suppose that $\limsup\limits_{k\to\infty} \dfrac{|a_k|}{b_k} = c < \infty$. By Proposition 1.98 there exists $N > 0$ such that $\dfrac{|a_k|}{b_k} < c+1$ for all $k \geqslant N$. This implies that $|a_k| < (c+1)b_k$ for all $k \geqslant N$; thus the convergence of $\sum\limits_{k=1}^{\infty} b_k$ and the comparison test imply that the series $\sum\limits_{k=1}^{\infty} |a_k|$ converges.

   (b) Suppose that $\liminf\limits_{k\to\infty} \dfrac{a_k}{b_k} = c > 0$. By Proposition 1.98 there exists $N > 0$ such that $\dfrac{a_k}{b_k} > \dfrac{c}{2}$ for all $k \geqslant N$. This implies that $a_k > \dfrac{c}{2} b_k$ for all $k \geqslant N$; thus the divergence of $\sum\limits_{k=1}^{\infty} b_k$ and the comparison test imply that the series $\sum\limits_{k=1}^{\infty} |a_k|$ diverges.

6. Let $\varepsilon > 0$ be given. Since $\{p_n\}_{n=1}^{\infty}$ is decreasing and $\lim\limits_{n\to\infty} p_n = 0$, there exists $N > 0$ such that
$$0 \leqslant p_n < \frac{\varepsilon}{2M+1} \qquad \text{whenever} \quad n \geqslant N\,.$$
Define $S_n = \sum\limits_{k=1}^{n} a_k$. Then if $n \geqslant N$ and $\ell \geqslant 0$,
$$\begin{aligned}\Big|\sum_{k=n}^{n+\ell} a_k p_k\Big| &= \big|a_n p_n + a_{n+1}p_{n+1} + a_{n+2}p_{n+2} + \cdots + a_{n+\ell-1}p_{n+\ell-1} + a_{n+\ell}p_{n+\ell}\big| \\ &= \big|(S_n - S_{n-1})p_n + (S_{n+1} - S_n)p_{n+1} + (S_{n+2} - S_{n+1})p_{n+2} + \cdots\cdots \\ &\quad + (S_{n+\ell-1} - S_{n+\ell-2})p_{n+\ell-1} + (S_{n+\ell} - S_{n+\ell-1})p_{n+\ell}\big| \\ &= \big|-S_{n-1}p_n + S_n(p_n - p_{n+1}) + S_{n+1}(p_{n+1} - p_{n+2}) + \cdots\cdots \\ &\quad + S_{n+\ell-1}(p_{n+\ell-1} - p_{n+\ell}) + S_{n+\ell}p_{n+\ell}\big| \\ &\leqslant |S_{n-1}p_n| + |S_n(p_n - p_{n+1})| + |S_{n+1}(p_{n+1} - p_{n+2})| + \cdots\cdots \\ &\quad + |S_{n+\ell-1}(p_{n+\ell-1} - p_{n+\ell})| + |S_{n+\ell}p_{n+\ell}| \\ &\leqslant Mp_n + M(p_n - p_{n+1}) + M(p_{n+1} - p_{n+2}) + \cdots\cdots \\ &\quad + M(p_{n+\ell-1} - p_{n+\ell}) + Mp_{n+\ell} \\ &= 2Mp_n < \frac{2M\varepsilon}{2M+1} < \varepsilon\,.\end{aligned}$$
The convergence of $\sum\limits_{k=1}^{\infty} a_k p_k$ follows from the Cauchy criterion (Theorem 2.65). □

Corollary 2.71. *1. The p-series* $\sum\limits_{k=1}^{\infty} \dfrac{1}{k^p}$ *converges if and only if* $p > 1$.

2. *The alternating series* $\sum\limits_{k=1}^{\infty} (-1)^k a_k$ *converges if* $\{a_k\}_{k=1}^{\infty}$ *is a decreasing convergent sequence with limit* 0.

Remark 2.72. It can be shown (and the proof is left as an exercise) that
$$\liminf_{k\to\infty} \frac{|x_{k+1}|}{|x_k|} \leqslant \liminf_{k\to\infty} \sqrt[k]{|x_k|} \leqslant \limsup_{k\to\infty} \sqrt[k]{|x_k|} \leqslant \limsup_{k\to\infty} \frac{|x_{k+1}|}{|x_k|}\,.$$
As a consequence, by the root test we obtain

1. if $\limsup\limits_{k\to\infty} \dfrac{|x_{k+1}|}{|x_k|} < 1$, the series $\sum\limits_{k=1}^{\infty} x_k$ converges absolutely, and
2. if $\liminf\limits_{k\to\infty} \dfrac{|x_{k+1}|}{|x_k|} > 1$, the series $\sum\limits_{k=1}^{\infty} x_k$ diverges.

This is called the ***ratio test***.

Example 2.73. The series $\sum\limits_{k=1}^{\infty} \dfrac{\sin k}{k^p}$ converges for $p > 0$ since

1. $\sum\limits_{k=1}^{n} \sin k = \dfrac{\cos\frac{1}{2} - \cos\frac{2k+1}{2}}{2\sin\frac{1}{2}}$; $\Big($thus $\Big|\sum\limits_{k=1}^{n} \sin k\Big| \leqslant \dfrac{1}{\sin\frac{1}{2}}\Big)$.

2. $\big\{\dfrac{1}{n^p}\big\}_{n=1}^{\infty}$ is decreasing and $\lim\limits_{n\to\infty} \dfrac{1}{n^p} = 0$.

We remark here that $\sum\limits_{k=1}^{\infty} \dfrac{\sin k}{k} = \dfrac{\pi - 1}{2}$. In fact, $\sum\limits_{k=1}^{\infty} \dfrac{\sin(kx)}{k}$ is the "Fourier series" of the function $\dfrac{\pi - x}{2}$.

Example 2.74. Let $\{x_k\}_{k=1}^{\infty}$ be a real sequence defined by

$$x_k = \begin{cases} 2^{-k} & \text{if } k \text{ is odd,} \\ 4^{-k} & \text{if } k \text{ is even,} \end{cases}$$

or $x_k = (3 + (-1)^k)^{-k}$. Then $\sqrt[k]{|x_k|} = (3 + (-1)^k)^{-1}$ which shows that

$$\liminf_{k\to\infty} \sqrt[k]{|x_k|} = \frac{1}{4} \qquad \text{and} \qquad \limsup_{k\to\infty} \sqrt[k]{|x_k|} = \frac{1}{2}.$$

Therefore, the root test implies that the series $\sum\limits_{k=1}^{\infty} x_k$ converges absolutely.

We can also compute the limit superior and limit inferior of $\dfrac{|x_{k+1}|}{|x_k|}$. Define

$$y_k = \frac{|x_{k+1}|}{|x_k|} = \frac{(3 + (-1)^{k+1})^{-k-1}}{(3 + (-1)^k)^{-k}} = \frac{1}{3 - (-1)^k}\Big(\frac{3 - (-1)^k}{3 + (-1)^k}\Big)^{-k}$$

and observe that $\lim\limits_{k\to\infty} y_{2k} = \infty$ and $\lim\limits_{k\to\infty} y_{2k+1} = 0$. Since $y_k \in [0, \infty)$, we conclude that 0 is the smallest cluster point of $\{y_k\}_{k=1}^{\infty}$ and ∞ is the largest "cluster point" of $\{y_k\}_{k=1}^{\infty}$. This shows that

$$\liminf_{k\to\infty} \frac{|x_{k+1}|}{|x_k|} = 0 \qquad \text{and} \qquad \limsup_{k\to\infty} \frac{|x_{k+1}|}{|x_k|} = \infty.$$

We note that in this case even if the series $\sum\limits_{k=1}^{\infty} x_k$ converges absolutely, $\limsup\limits_{k\to\infty} \dfrac{|x_{k+1}|}{|x_k|} > 1$. Therefore, the condition $\limsup\limits_{k\to\infty} \dfrac{|x_{k+1}|}{|x_k|} > 1$ cannot be used to guarantee the divergence of the series $\sum\limits_{k=1}^{\infty} x_k$.

2.5 Exercises

§ 2.1 Euclidean Spaces and Vector Spaces

Problem 2.1. Let $\|\cdot\|: \mathbb{F}^n \to \mathbb{R}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, be defined by

$$\|\boldsymbol{x}\|_p \equiv \begin{cases} \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} & \text{if } 1 \leqslant p < \infty\,, \\ \max\big\{|x_1|, \cdots, |x_n|\big\} & \text{if } p = \infty\,, \end{cases} \qquad \boldsymbol{x} = (x_1, \cdots, x_n)\,.$$

Complete the following.

1. Prove the Hölder inequality $\big|\langle \boldsymbol{x}, \boldsymbol{y}\rangle\big| \leqslant \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$, where $p, q \in [1, \infty]$ satisfy $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

2. Show that $\|\cdot\|_p$ is indeed a norm on $\mathbb{F}^n$ for all $1 \leqslant p \leqslant \infty$.

3. Show that $\|\boldsymbol{x}\|_\infty = \lim\limits_{p\to\infty} \|\boldsymbol{x}\|_p$ for all $\boldsymbol{x} \in \mathbb{F}^n$.

4. Show that for each $1 \leqslant p, q \leqslant \infty$ and $p \neq q$, $\|\cdot\|_p$ and $\|\cdot\|_q$ are equivalent norms.

Hint: 1. Prove first the Young inequality (if you do not know this inequality)

$$ab \leqslant \frac{1}{p}a^p + \frac{1}{q}b^q \qquad \forall\, a, b \geqslant 0 \text{ and } p, q \in (1, \infty) \text{ satisfying } \frac{1}{p} + \frac{1}{q} = 1\,,$$

Problem 2.2. Complete the following.

1. For $f \in \mathscr{C}([a, b]; \mathbb{R})$, define

$$\|f\|_p = \begin{cases} \Big(\int_a^b |f(x)|^p\, dx\Big)^{\frac{1}{p}} & \text{if } 1 \leqslant p < \infty\,, \\ \max\limits_{x\in[a,b]} \big|f(x)\big| & \text{if } p = \infty\,. \end{cases}$$

   Show that $\|\cdot\|_p$ is a norm on $\mathscr{C}([a, b]; \mathbb{R})$.

2. Show that $\|f\|_\infty = \lim\limits_{p\to\infty} \|f\|_p$ for all $f \in \mathscr{C}([a, b]; \mathbb{R})$.

3. Are $\|\cdot\|_p$ and $\|\cdot\|_q$ equivalent norms on $\mathscr{C}([a, b]; \mathbb{R})$ for any $1 \leqslant p, q \leqslant \infty$?

Problem 2.3. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Show that for each $\boldsymbol{x} = (x_1, \cdots, x_n)$ and $\boldsymbol{y} = (y_1, \cdots, y_n)$ in $\mathbb{F}^n$,

$$\|\boldsymbol{x}\|_1 = \sup_{\|\boldsymbol{y}\|_\infty = 1} \Big| \sum_{k=1}^n x_k y_k \Big| = \sup \Big\{ \Big| \sum_{k=1}^n x_k y_k \Big| \, \Big| \, \|\boldsymbol{y}\|_\infty = 1 \Big\}$$

and

$$\|\boldsymbol{y}\|_\infty = \sup_{\|\boldsymbol{x}\|_1 = 1} \Big| \sum_{k=1}^n x_k y_k \Big| = \sup \Big\{ \Big| \sum_{k=1}^n x_k y_k \Big| \, \Big| \, \|\boldsymbol{x}\|_1 = 1 \Big\} .$$

Problem 2.4. Let $\mathcal{V}$ be a vector space (over a scalar field $\mathbb{F}$), and A, B be subset of $\mathcal{V}$. Show that

$$A + B = \bigcup_{\boldsymbol{a} \in A} (\boldsymbol{a} + B) = \bigcup_{\boldsymbol{b} \in B} (\boldsymbol{b} + A) .$$

Problem 2.5. A set A in a vector space $\mathcal{V}$ is called ***convex*** if for all $x, y \in A$, the line segment joining x and y, denoted by $\overline{xy}$, lies in A.

1. Show that open r-balls and closed r-balls in any normed vector spaces are convex.

2. Show that an ellipsoid $E \equiv \big\{ \boldsymbol{x} \in \mathbb{R}^n \,\big|\, \boldsymbol{x}^{\mathrm{T}} Q \boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}} \boldsymbol{x} + \boldsymbol{c} \leqslant 0 \big\}$ is convex, where Q is a positive semi-definite $n \times n$ matrix and $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$ are vectors.

3. Show that a polytope $P \equiv \big\{ \boldsymbol{x} \in \mathbb{R}^n \,\big|\, A\boldsymbol{x} \leqslant \boldsymbol{b} \big\}$ is convex, where A is an $m \times n$ matrix, $\boldsymbol{b} \in \mathbb{R}^m$ is a vector, and $\leqslant$ is defined by $\boldsymbol{c} \leqslant \boldsymbol{d}$ if and only if $c_i \leqslant d_i$ for all components.

4. Show that if C_α is convex for all $\alpha \in I$, then $\bigcap\limits_{\alpha \in I} C_\alpha$ is convex.

5. Show that if $C_1, C_2, \cdots, C_N$ are convex sets and $\mu_1, \mu_2, \cdots, \mu_N$ are real numbers, then

   $$\mu_1 C_1 + \mu_2 C_2 + \cdots + \mu_N C_N \equiv \big\{ \mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_N x_n \,\big|\, x_k \in C_k \text{ for } 1 \leqslant k \leqslant N \big\}$$

   is convex.

6. Show that if $C_1, C_2, \cdots, C_N$ are convex sets, then the Cartesian product $C_1 \times C_2 \times \cdots \times C_N$ is convex.

7. Let A be an $m \times n$ matrix, and C be a convex set in $\mathbb{R}^n$, D be a convex set in $\mathbb{R}^m$. Show that

   $$A(C) \equiv \big\{ A\boldsymbol{x} \,\big|\, \boldsymbol{x} \in C \big\} \qquad \text{and} \qquad A^{-1}(D) \equiv \big\{ \boldsymbol{x} \,\big|\, A\boldsymbol{x} \in D \big\}$$

   are convex.

8. Let $\Delta_k \equiv \Big\{(\lambda_1, \cdots, \lambda_k) \in \mathbb{R}^k \,\Big|\, 0 \leqslant \lambda_i \leqslant 1 \text{ for all } 1 \leqslant i \leqslant k \text{ and } \sum\limits_{i=1}^{k} \lambda_i = 1\Big\}$. A convex combination of k vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k \in \mathbb{R}^n$ is a sum of the form $\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k$ for some $(\lambda_1, \cdots, \lambda_k) \in \Delta_k$. Show that the collection of all convex combinations of k given vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ is convex; that is, show that

$$\big\{\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k \,\big|\, (\lambda_1, \lambda_2, \cdots, \lambda_k) \in \Delta_k\big\}$$

is convex.

9. Let S be a subset in $\mathbb{R}^n$. Show that the collection of all convex combinations of finitely many vectors from S is convex; that is, show that

$$\big\{\lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k \,\big|\, k \in \mathbb{N}\,, \boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in S \text{ and } (\lambda_1, \cdots, \lambda_k) \in \Delta_k\big\}$$

is convex. The set defined above is called the convex full of S and is sometimes denoted by $\operatorname{conv}(S)$.

§ 2.2 Normed Vector Spaces, Inner Product Spaces and Metric Spaces

Problem 2.6. Show that

$$\|A\|_1 = \max\left\{\sum_{i=1}^{n} |a_{i1}|, \sum_{i=1}^{n} |a_{i2}|, \cdots, \sum_{i=1}^{n} |a_{im}|\right\} \qquad \forall\, A \in \mathcal{M}_{n\times m}\,.$$

Hint: Mimic the computation of $\|A\|_\infty$ in Example 2.19, or make use of Exercise Problem 1.15 and 2.3.

Problem 2.7. Let $\mathcal{M}_{n\times m}(\mathbb{F})$ be collection of $n \times m$ matrices with entries in $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. For $A \in \mathcal{M}_{n\times m}(\mathbb{F})$, define

$$\|A\|_p = \sup_{\|\boldsymbol{x}\|_p = 1} \|A\boldsymbol{x}\|_p = \sup_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|A\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p}\,.$$

1. Show that $\|\cdot\|_p$ is a norm on $\mathcal{M}_{n\times m}(\mathbb{F})$.

2. Show that $\|A\|_2 = \sqrt{\text{the maximum eigenvalue of } A^\dagger A}$, where $A^\dagger$ is the conjugate transpose of A.

3. Show that $\|A\|_\infty = \max\Big\{\sum\limits_{k=1}^{m} |a_{1k}|, \sum\limits_{k=1}^{m} |a_{2k}|, \cdots, \sum\limits_{k=1}^{m} |a_{nk}|\Big\}$ if $A \in \mathcal{M}_{n\times m}(\mathbb{F})$.

4. Show that $\|A\|_1 = \max\Big\{\sum\limits_{k=1}^n |a_{k1}|, \sum\limits_{k=1}^n |a_{k2}|, \cdots, \sum\limits_{k=1}^n |a_{km}|\Big\}$ if $A \in \mathcal{M}_{n\times m}(\mathbb{F})$.

5. Show that $\|A\|_2^2 \leqslant \|A\|_1 \|A\|_\infty$ for all $A \in \mathcal{M}_{n\times m}(\mathbb{F})$.

Problem 2.8. Let $\mathcal{M}_{n\times m}(\mathbb{F})$ be the collection of all $n \times m$ matrices with entries in $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Define a function $\|\cdot\|_{p,q} : \mathcal{M}_{n\times m}(\mathbb{F}) \to \mathbb{R}$ by

$$\|A\|_{p,q} = \sup_{\|\boldsymbol{x}\|_p = 1} \|A\boldsymbol{x}\|_q\,,$$

here we recall that $\|\cdot\|_p$ is the p-norm on $\mathbb{F}^n$ given in Problem 2.1. If $p = q$, we simply use $\|A\|_p$ to denote $\|A\|_{p,q}$. Complete the following.

1. Show that $\|A\|_{p,q} = \sup\limits_{\boldsymbol{x}\neq\boldsymbol{0}} \dfrac{\|A\boldsymbol{x}\|_q}{\|\boldsymbol{x}\|_p}$ for all $p, q \geqslant 1$.

2. Show that $\|A\|_{p,q} = \inf\big\{M \in \mathbb{F} \,\big|\, \|A\boldsymbol{x}\|_q \leqslant M\|\boldsymbol{x}\|_p\ \forall\, \boldsymbol{x} \in \mathbb{F}^m\big\}$.

3. $\|A\boldsymbol{x}\|_q \leqslant \|A\|_{p,q}\|\boldsymbol{x}\|_p$ for all $\boldsymbol{x} \in \mathbb{F}^m$.

4. $\|\cdot\|_{p,q}$ defines a norm on $\mathcal{M}_{n\times m}(\mathbb{F})$.

5. Let $\{A_k\}_{k=1}^\infty \subseteq \mathcal{M}_{n\times m}(\mathbb{F})$. Show that $\lim\limits_{k\to\infty} \|A_k\|_{p,q} = 0$ if and only if each entry of A_k converges to 0. In other words, by writing $A_k = \big[a_{ij}^{(k)}\big]_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$, show that $\lim\limits_{k\to\infty} \|A_k\|_{p,q} = 0$ if and only if $\lim\limits_{k\to\infty} a_{ij}^{(k)} = 0$ for all $1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$. In particular, $A_k \to A$ in the sense that $\|A_k - A\|_{p,q} \to 0$ as $k \to \infty$ if and only if the (i,j)-th entry of A_k converges to (i,j)-th entry of A for all $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$.

Problem 2.9. Let $\mathcal{M}_{n\times m}(\mathbb{F})$ be the collection of all $n \times m$ matrices with entries in $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Define $\|\cdot\|_F : \mathcal{M}_{n\times m}(\mathbb{F}) \to \mathbb{R}$ by

$$\|A\|_F \equiv \Big(\sum_{i=1}^n \sum_{j=1}^m |a_{ij}|^2\Big)^{\frac{1}{2}}.$$

1. Show that $\|A\|_F^2 = \operatorname{tr}(A^\dagger A)$, where $A^\dagger$ is the conjugate transpose of A, and $\operatorname{tr}(M)$ is the trace of square matrix M.

2. Show that $\|\cdot\|_F$ is a norm on $\mathcal{M}_{n\times m}(\mathbb{F})$ (for all $n, m \in \mathbb{N}$). This norm is called the Frobenius norm of matrices.

3. Show that $\|AB\|_F \leqslant \|A\|_F\|B\|_F$ whenever $A \in \mathcal{M}_{n\times m}(\mathbb{F})$ and $B \in \mathcal{M}_{m\times p}(\mathbb{F})$.

Hint: 3. Let $A = \big[\boldsymbol{a}_1 \vdots \boldsymbol{a}_2 \vdots \cdots \vdots \boldsymbol{a}_m\big]$ and $B = \big[\boldsymbol{b}_1 \vdots \boldsymbol{b}_2 \vdots \cdots \vdots \boldsymbol{b}_m\big]^{\mathrm{T}}$; that is, $\boldsymbol{a}_k$ is the k-th column of A and $\boldsymbol{b}_\ell$ is the ℓ-th row of B. Then $AB = \sum\limits_{k=1}^{m} \boldsymbol{a}_k \boldsymbol{b}_k$. First show that $\|\boldsymbol{a}_k \boldsymbol{b}_k^{\mathrm{T}}\|_F = \|\boldsymbol{a}_k\|_2 \|\boldsymbol{b}_k\|_2$ and use the triangle inequality to conclude the desired equality.

Problem 2.10. Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$. Show that

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0 \qquad \text{if and only if} \qquad \|\boldsymbol{y}\| \leqslant \|\lambda \boldsymbol{x} + \boldsymbol{y}\| \quad \forall\, \lambda \in \mathbb{F}\,,$$

where $\|\cdot\|$ is the norm induced by the inner product $\langle \cdot, \cdot \rangle$.

Problem 2.11. Let $(\mathcal{V}, +, \cdot, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{R}$, and define $\|\boldsymbol{v}\| = \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{1/2}$ for all $\boldsymbol{v} \in \mathcal{V}$. Show that

1. $2\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{y}\|^2 = \|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2$ (parallelogram law).
2. $\big|\|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\big| \leqslant \|\boldsymbol{x}+\boldsymbol{y}\|\|\boldsymbol{x}-\boldsymbol{y}\| \leqslant \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$.
3. $4\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2$ (polarization identity).

Can the p-norm $\|\cdot\|_p$ on $\mathbb{R}^n$ be induced from any inner product (on $\mathbb{R}^n$) for $p \neq 2$?

Problem 2.12. Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{C}$. Show the polarization identity

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \frac{1}{4}\Big(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + i\|\boldsymbol{x}+i\boldsymbol{y}\|^2 - i\|\boldsymbol{x}-i\boldsymbol{y}\|^2\Big) \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}\,.$$

Problem 2.13. Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ be three normed vector spaces such that $X, Y \subseteq Z$ and

$$\|\boldsymbol{x}\|_Z \leqslant C\|\boldsymbol{x}\|_X \quad \forall\, \boldsymbol{x} \in X \qquad \text{and} \qquad \|\boldsymbol{y}\|_Z \leqslant C\|\boldsymbol{y}\|_Y \quad \forall\, \boldsymbol{y} \in Y\,.$$

Define

$$E = \big\{\boldsymbol{a} \in Z \,\big|\, \|\boldsymbol{a}\|_E \equiv \max\{\|\boldsymbol{a}\|_X, \|\boldsymbol{a}\|_Y\} < \infty\big\}$$

and

$$F = \big\{\boldsymbol{a} \in Z \,\big|\, \|\boldsymbol{a}\|_F \equiv \inf_{\substack{\boldsymbol{a}=\boldsymbol{x}+\boldsymbol{y} \\ \boldsymbol{x}\in X, \boldsymbol{y}\in Y}} (\|\boldsymbol{x}\|_X + \|\boldsymbol{y}\|_Y) < \infty\big\}\,.$$

1. Show that $(E, \|\cdot\|_E)$ is a normed vector space, and $E = X \cap Y$.

2. Show that $(F, \|\cdot\|_F)$ is a normed vector space. The space F is usually denoted by $X+Y$.

Problem 2.14. Let (M,d) be a metric space. Define $\rho : M \times M \to \mathbb{R}$ by

$$\rho(x,y) = \frac{d(x,y)}{1+d(x,y)}\,.$$

Show that (M,ρ) is also a metric space.

Problem 2.15. Let $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$d(\boldsymbol{x},\boldsymbol{y}) = \begin{cases} |x_1-y_1| & \text{if } x_2 = y_2\,, \\ |x_1-y_1|+|x_2-y_2|+1 & \text{if } x_2 \neq y_2\,, \end{cases} \quad \text{where } \boldsymbol{x} = (x_1,x_2) \text{ and } \boldsymbol{y} = (y_1,y_2).$$

Show that d is a metric on $\mathbb{R}^2$.

§ 2.3 Sequences in Metric Spaces

Problem 2.16. Let (M,d) be a metric space, and $\{x_n\}_{n=1}^{\infty}$ be a sequence in M. Show that $\{x_n\}_{n=1}^{\infty}$ is Cauchy if and only if

$$\forall\, \varepsilon > 0, \exists\, y \in M \ni \#\big\{n \in \mathbb{N} \,\big|\, x_n \notin B(y,\varepsilon)\big\} < \infty\,.$$

Problem 2.17. Show that

1. Every convergent sequence in a metric space is a Cauchy sequence.

2. If a subsequence of a Cauchy sequence converges to x, then the sequence converges to x.

3. x is a cluster point of $\{x_k\}_{k=1}^{\infty}$ if and only if $\forall\, \varepsilon > 0$ and $N > 0$, $\exists\, k > N$ with $d(x_k, x) < \varepsilon$.

4. x is a cluster point of $\{x_k\}_{k=1}^{\infty}$ if and only if there is a subsequence converging to x.

5. $x_k \to x$ as $k \to \infty$ if and only if every subsequence of $\{x_k\}_{k=1}^{\infty}$ converges to x.

6. $x_k \to x$ as $k \to \infty$ if and only if every proper subsequence of $\{x_k\}_{k=1}^{\infty}$ has a further subsequence that converges to x.

Problem 2.18. Let $(\mathcal{V}, \|\cdot\|)$ be a norm space, $\{a_k\}_{k=1}^{\infty}$ be a sequence in $\mathcal{V}$, and $\{b_n\}_{n=1}^{\infty}$ be the Cesàro mean of $\{a_k\}_{k=1}^{\infty}$; that is, $b_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} a_k$.

1. Show that if $\{a_k\}_{k=1}^{\infty}$ converges to a, then $\{b_n\}_{n=1}^{\infty}$ converges to a.

2. Is the converse statement "if the Cesàro mean $\{b_n\}_{n=1}^{\infty}$ of a sequence $\{a_k\}_{k=1}^{\infty}$ converges, then the sequence $\{a_k\}_{k=1}^{\infty}$ also converges" true?

Problem 2.19. Show that $(\mathbb{C}, |\cdot|)$ is complete.

Problem 2.20. Let A be a subset in $\mathbb{R}^n$. Show that A is bounded if and only if every sequence in A has a convergent subsequence.

Problem 2.21. Let $\mathcal{V}$ be a vector space over field $\mathbb{F}$, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ be a basis of $\mathcal{V}$; that is, every vector $\mathbf{e} \in \mathcal{V}$ can be expressed (uniquely) by $\boldsymbol{x} = \sum\limits_{i=1}^{n} x_i \mathbf{e}_i$ with $x_i \in \mathbb{F}$ for all $1 \leqslant i \leqslant n$. From Example 2.28 we know that $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{F}$ defined by

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{j=1}^{n} x_i \overline{y_i} \qquad \text{if } \boldsymbol{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j \text{ and } \boldsymbol{y} = \sum_{j=1}^{n} y_j \mathbf{e}_j .$$

is an inner product. Show that $(\mathcal{V}, \|\cdot\|_2)$ is a Banach space, where $\|\cdot\|_2$ is the norm induced by the inner product.

Problem 2.22. Let (M, d) be a metric space. Two Cauchy sequences $\{p_n\}_{n=1}^{\infty}$ and $\{q_n\}_{n=1}^{\infty}$ in M are said to be equivalent, denoted by $\{p_n\}_{n=1}^{\infty} \sim \{q_n\}_{n=1}^{\infty}$, if

$$\lim_{n \to \infty} d(p_n, q_n) = 0 .$$

1. Prove that $\sim$ is an equivalence relation; that is, show that

   (a) $\{p_n\}_{n=1}^{\infty} \sim \{p_n\}_{n=1}^{\infty}$.

   (b) If $\{p_n\}_{n=1}^{\infty} \sim \{q_n\}_{n=1}^{\infty}$, then $\{q_n\}_{n=1}^{\infty} \sim \{p_n\}_{n=1}^{\infty}$.

   (c) If $\{p_n\}_{n=1}^{\infty} \sim \{q_n\}_{n=1}^{\infty}$ and $\{q_n\}_{n=1}^{\infty} \sim \{r_n\}_{n=1}^{\infty}$, then $\{p_n\}_{n=1}^{\infty} \sim \{r_n\}_{n=1}^{\infty}$.

2. Let $\{p_n\}_{n=1}^{\infty}$ and $\{q_n\}_{n=1}^{\infty}$ be two Cauchy sequences. Show that the sequence $\big\{d(p_n, q_n)\big\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$; thus is convergent.

3. Let M^* be the set of all equivalence classes. If $P, Q \in M^*$, we define
$$d^*(P,Q) = \lim_{n\to\infty} d(p_n, q_n)\,,$$
where $\{p_n\}_{n=1}^\infty \in P$ and $\{q_n\}_{n=1}^\infty \in Q$. Show that the definition above is well-defined; that is, show that if $\{p_n'\}_{n=1}^\infty \in P$ and $\{q_n'\}_{n=1}^\infty \in Q$ are another two Cauchy sequences, then $\lim\limits_{n\to\infty} d(p_n, q_n) = \lim\limits_{n\to\infty} d(p_n', q_n')$.

4. Define $\varphi : M \to M^*$ as follows: for each $x \in M$, $\{x_n\}_{n=1}^\infty$, where $x_n \equiv x$ for all $n \in \mathbb{N}$, is a Cauchy sequence in M. Then $\{x_n\}_{n=1}^\infty \in \varphi(x)$ for one particular $\varphi(x) \in M^*$. In other words, $\varphi(x)$ is the equivalence class where $\{x_n\}_{n=1}^\infty$ belongs to. Show that
$$d^*\big(\varphi(x), \varphi(y)\big) = d(x,y) \qquad \forall\, x, y \in M.$$

5. Show that $\varphi(M)$ is dense in M^*.

6. Show that (M^*, d^*) is a complete metric space. The metric space (M^*, d^*) is called the completion of (M, d).

§ 2.4 Series of Real Numbers and Vectors

Problem 2.23. Let $\{a_n\}_{n=1}^\infty$ and $\{x_n\}_{n=1}^\infty$ be two sequences of real numbers, and $|x_n - x_{n+1}| < a_n$ for all $n \in \mathbb{N}$. Show that $\{x_n\}_{n=1}^\infty$ converges if $\sum\limits_{n=1}^\infty a_n$ converges.

Problem 2.24. Let $\sum\limits_{k=1}^\infty a_k$ be a conditionally convergent series. Show that $\sum\limits_{k=1}^\infty \big[1+\text{sgn}(a_k)\big]a_k$ and $\sum\limits_{k=1}^\infty \big[1-\text{sgn}(a_k)\big]a_k$ both diverge. Here the sign function sgn is defined by
$$\text{sgn}(a) = \begin{cases} 1 & \text{if } a > 0\,, \\ 0 & \text{if } a = 0\,, \\ -1 & \text{if } a < 0\,. \end{cases}$$

Problem 2.25. Let $\{a_k\}_{k=1}^\infty \subseteq \mathbb{R}$ ba a sequence. A series $\sum\limits_{k=1}^\infty b_k$ is said to be a rearrangement of the series $\sum\limits_{k=1}^\infty a_k$ if there exists a rearrangement π of $\mathbb{N}$; that is, $\pi : \mathbb{N} \to \mathbb{N}$ is bijective, such that $b_k = a_{\pi(k)}$.

1. Show that if $\sum\limits_{k=1}^\infty a_k$ converges absolutely, then any rearrangement of the series $\sum\limits_{k=1}^\infty a_k$ converges and has the value $\sum\limits_{k=1}^\infty a_k$.

2. Show that if $\sum\limits_{k=1}^{\infty} a_k$ is conditionally convergent, then for each $r \in \mathbb{R}$, there exists a rearrangement $\sum\limits_{k=1}^{\infty} a_{\pi(k)}$ of the series $\sum\limits_{k=1}^{\infty} a_k$ such that $\sum\limits_{k=1}^{\infty} a_{\pi(k)} = r$.

Problem 2.26. Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be sequences of real numbers such that $a_n, b_n > 0$ for all $n \geqslant N$. Define

$$c_n = b_n - b_{n+1}\frac{a_{n+1}}{a_n} \qquad \forall\, n \in \mathbb{N}\,. \tag{$\star$}$$

1. Show that if there exists a constant $r > 0$ such that $r < c_n$ for all $n \geqslant N$, then $\sum\limits_{k=1}^{\infty} a_k$ converges.

   Hint: Rewrite ($\star$) as $b_n = c_n + \dfrac{a_{n+1}}{a_n} b_{n+1}$ and then obtain

$$\begin{aligned} b_N &= c_N + \frac{a_{N+1}}{a_N} b_{N+1} = c_N + \frac{a_{N+1}}{a_N}\Big(c_{N+1} + \frac{a_{N+2}}{a_{N+1}} b_{N+2}\Big) = c_N + \frac{a_{N+1}}{a_N} c_{N+1} + \frac{a_{N+2}}{a_N} b_{N+2} \\ &= c_N + \frac{a_{N+1}}{a_N} c_{N+1} + \frac{a_{N+2}}{a_N}\Big(c_{N+2} + \frac{a_{N+3}}{a_{N+2}} b_{N+3}\Big) = \cdots \\ &= c_N + \frac{a_{N+1}}{a_N} c_{N+1} + \frac{a_{N+2}}{a_N} c_{N+2} + \cdots + \frac{a_{N+n}}{a_N} c_{N+n} + \frac{a_{N+n+1}}{a_N} b_{N+n+1}\,. \end{aligned}$$

   Use the fact that $0 < r < c_n$ for all $n \geqslant N$ to conclude that

$$\sum_{k=N}^{N+n} a_k \leqslant \frac{a_N b_N}{r} \qquad \forall\, n \in \mathbb{N}\,.$$

   Note that then the sequence of partial sum of $\sum\limits_{k=1}^{\infty} a_k$ then is bounded from above $\Big($by $\sum\limits_{k=1}^{N-1} a_k + \dfrac{a_N b_N}{r}\Big)$.

2. Show that if $\sum\limits_{k=1}^{\infty} \dfrac{1}{b_k}$ diverges and $c_n \leqslant 0$ for all $n \geqslant N$, then $\sum\limits_{k=1}^{\infty} a_k$ diverges.

   Hint: The fact that $c_n \leqslant 0$ for all $n \geqslant N$ implies that $b_n a_n \leqslant b_{n+1} a_{n+1}$ for all $n \geqslant N$. Use this fact to conclude that

$$\frac{a_N b_N}{b_n} \leqslant a_n \qquad \forall\, n \geqslant N$$

   and then apply the direct comparison test to conclude that $\sum\limits_{k=1}^{\infty} a_k$ diverges.

Problem 2.27. Let $\sum\limits_{k=1}^{\infty} a_k$ be a series with positive terms, and $\lim\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n} = 1$. We know that the ratio test fails when this happens, but there are some refined results concerning this particular case.

1. **(Raabe's test)**:
   (a) If there exists a constant $\mu > 1$ such that $\dfrac{a_{n+1}}{a_n} < 1 - \dfrac{\mu}{n}$ for all $n \geqslant N$, then $\sum\limits_{k=1}^{\infty} a_k$ converges.
   (b) If there exists a constant $0 < \mu < 1$ such that $\dfrac{a_{n+1}}{a_n} > 1 - \dfrac{\mu}{n}$ for all $n \geqslant N$, then $\sum\limits_{k=1}^{\infty} a_k$ diverges.
2. **(Gauss's test)**: Suppose that there exist a positive constant $\epsilon > 0$, a constant μ, and a **bounded** sequence $\{R_n\}_{n=1}^{\infty}$ such that
$$\frac{a_{n+1}}{a_n} = 1 - \frac{\mu}{n} + \frac{R_n}{n^{1+\epsilon}} \qquad \text{for all } n \geqslant N\,.$$
   (a) If $\mu > 1$, then $\sum\limits_{k=1}^{\infty} a_k$ converges. (b) If $\mu \leqslant 1$, then $\sum\limits_{k=1}^{\infty} a_k$ diverges.

Problem 2.28. Let a_n be defined by $a_n = \begin{cases} \dfrac{n+1}{2^n} & \text{if } n \text{ is odd}\,, \\ \dfrac{n}{3^n} & \text{if } n \text{ is even}\,. \end{cases}$ Compute the value of $\liminf\limits_{n\to\infty} \sqrt[n]{a_n}$, $\limsup\limits_{n\to\infty} \sqrt[n]{a_n}$, $\liminf\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n}$ and $\limsup\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n}$, and conclude that whether the series $\sum\limits_{n=1}^{\infty} a_n$ is convergent or not.
Hint: You can use $\lim\limits_{n\to\infty} \sqrt[n]{n} = \lim\limits_{n\to\infty} \sqrt[n]{n+1} = 1$ without proving it.

Problem 2.29. Let $\alpha \in \mathbb{R}$, $\alpha > \dfrac{1}{3}$. Discuss the absolute convergence or the conditional convergence of the series $\sum\limits_{k=2}^{\infty} \dfrac{(-1)^k}{k^\alpha + (-1)^k}$.

Problem 2.30. Determine whether the following series converge or not. Also test for their absolute convergence.

1. $\sum\limits_{n=1}^{\infty} \sin(n^{-\alpha})$, $\alpha > 0$;
2. $\sum\limits_{n=1}^{\infty} \dfrac{\log(n+1) - \log n}{\arctan \frac{2}{n}}$;
3. $\sum\limits_{n=1}^{\infty} \dfrac{a(a+1)\cdots(a+n-1)b(b+1)\cdots(b+n-1)}{1\cdot 2\cdots n\cdot c(c+1)\cdots(c+n-1)}$;

4. $\sum\limits_{n=1}^{\infty} \frac{(-1)^n}{n+1}\left(1+\frac{1}{3}+\cdots+\frac{1}{2n+1}\right)$;

5. $\sum\limits_{n=2}^{\infty} \frac{(-1)^n}{\sqrt{n}+(-1)^n}$;

The a, b, c in (3) are real number except negative integers.

Problem 2.31. Complete the following.

1. Show that $\sum\limits_{k=1}^{\infty}\left(1-\frac{1}{\sqrt{k}}\right)^k$ converges.

2. Show that $\sum\limits_{k=2}^{\infty}\frac{\log(k+1)-\log k}{(\log k)^2}$ converges.

3. Use Gauss's test to show that both the general harmonic series $\sum\limits_{k=1}^{\infty}\frac{1}{ak+b}$, where $a \neq 0$, and the series $\sum\limits_{k=1}^{\infty}\frac{1}{\sqrt{k}}$ diverge.

4. Show that $\sum\limits_{k=1}^{\infty}\frac{k!}{(\alpha+1)(\alpha+2)\cdots(\alpha+k)}$ converges if $\alpha > 1$ and diverges if $\alpha \leqslant 1$.

5. Test the following "hypergeometric" series for convergence or divergence:

   (a) $\sum\limits_{k=1}^{\infty}\frac{\alpha(\alpha+1)(\alpha+2)\cdots(\alpha+k-1)}{\beta(\beta+1)(\beta+2)\cdots(\beta+k-1)} = \frac{\alpha}{\beta}+\frac{\alpha(\alpha+1)}{\beta(\beta+1)}+\frac{\alpha(\alpha+1)(\alpha+2)}{\beta(\beta+1)(\beta+2)}+\cdots$.

   (b) $1+\frac{\alpha\cdot\beta}{1\cdot\gamma}+\frac{\alpha(\alpha+1)\cdot\beta(\beta+1)}{1\cdot 2\gamma\cdot(\gamma+1)}+\frac{\alpha(\alpha+1)(\alpha+2)\cdot\beta(\beta+1)(\beta+2)}{1\cdot 2\cdot 3\cdot\gamma(\gamma+1)(\gamma+2)}+\cdots$.

Problem 2.32. Consider the function $f(x) = \sum\limits_{k=1}^{\infty}\frac{\sin(kx)}{k}$.

1. Find the domain of f.

2. Show that for each $\varepsilon > 0$ and $0 < \delta < \pi$, there exists $N > 0$ and N depends only on ε and δ but is independent of x, such that

$$\Big|\sum_{k=n}^{n+p}\frac{\sin(kx)}{k}\Big| < \varepsilon \qquad \forall\, n \geqslant N, p \geqslant 0 \text{ and } x \in [\delta, 2\pi-\delta]\,.$$

§ 2.5 Chapter Review

Problem 2.33 (True or False). Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. Given two sets A and B. Then $A \times B$ is countable if and only if A and B are countable.

2. Let $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ be a sequence and $\limsup\limits_{n\to\infty} x_n = x$. Then $\sup\limits_{n\in\mathbb{N}} x_n = x$.

3. The set $\{(x,y) \in \mathbb{R}^2 \,|\, x+y \in \mathbb{Q}\}$ is countable.

4. Let $\{x_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ be a sequence such that $|x_n - x_{n+1}| \leqslant \dfrac{1}{n}$. Then $\{x_n\}_{n=1}^{\infty}$ converges in $\mathbb{R}$.

5. If a bounded sequence $\{x_n\}_{n=1}^{\infty}$ in $\mathbb{R}$ satisfies $x_{n+1} - \epsilon_n \leqslant x_n$ for $n \in \mathbb{N}$, where $\sum\limits_{n=1}^{\infty} \epsilon_n$ is an absolute convergent series; that is, the partial sum $\sum\limits_{n=1}^{k} |\epsilon_n|$ converges as $k \to \infty$, then $\{x_n\}_{n=1}^{\infty}$ converges.

6. Let $\pi : \mathbb{N} \to \mathbb{N}$ be one-to-one and onto (such map π is called a rearrangement), and $\{x_n\}_{n=1}^{\infty}$ is a convergent sequence. Then $\{x_{\pi(n)}\}_{n=1}^{\infty}$ is also convergent.

7. Let $A \subseteq \mathbb{R}$ satisfy
$$\sup\Big\{\sum_{b\in B} |b| \,\Big|\, B \text{ is a non-empty finite subsets of } A\Big\} < \infty\,.$$
Then $\{x \in A \,|\, x \neq 0\}$ is countable.

8. Any rearrangement of the series $\sum\limits_{n=1}^{\infty} x_n$ diverges if and only if x_n does not tend to 0 as $n \to \infty$.

9. If $\{x_n\}_{n=1}^{\infty}$ is a sequence of distinct non-zero real numbers such that $\lim\limits_{n\to\infty} x_n = 0$, then the set $\{mx_n \,|\, m \in \mathbb{Z}, n \in \mathbb{N}\}$ is dense in $\mathbb{R}$.

Chapter 3

Elementary Point-Set Topology

3.1 Limit Points and Interior Points of Sets

Definition 3.1. Let (M, d) be a metric space, and A be a subset of M.

1. A point $x \in M$ is called a ***limit point*** of A if there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in A such that $\{x_n\}_{n=1}^{\infty}$ converges to x.

2. The ***closure*** of A is the collection of all limit points of A is denoted by $\bar{A}$ or $\text{cl}(A)$.

3. A point $x \in M$ is called an ***interior point*** of A if there exists $r > 0$ such that the r-ball about x is contained in A; that is, $B(x, r) \subseteq A$.

4. The ***interior*** of A is the collection of all interior points of A and is denoted by $\mathring{A}$ or $\text{int}(A)$.

5. A point $x \in M$ is called an ***exterior point*** of A if x is an interior point of $A^{\complement}$, and the collection of all exterior points of A is called the ***exterior*** of A.

Example 3.2. For $a, b \in \mathbb{R}$ and $a < b$, consider the interval I in $(R, |\cdot|)$ with end-points a and b. Then $\bar{\mathrm{I}} = [a, b]$ and $\mathring{\mathrm{I}} = (a, b)$.

Remark 3.3. By the definition of the convergence of sequences in metric spaces, we have the following equivalent definition: A point $x \in M$ is called a limit point of A if for every $\varepsilon > 0$, $B(x, \varepsilon)$ contains points in A; that is, $\forall\, \varepsilon > 0, B(x, \varepsilon) \cap A \neq \varnothing$.

Remark 3.4. 1. If $x \in A$, then x is a limit point of A. In other words, $A \subseteq \bar{A}$.

2. If $x \in \mathring{A}$, then $x \in A$; thus $\mathring{A} \subseteq A$.

Theorem 3.5. *Let (M,d) be a metric space, and A, B be subsets of M. If $A \subseteq B$, then $\bar{A} \subseteq \bar{B}$ and $\mathring{A} \subseteq \mathring{B}$.*

Proof. 1. Let $x \in \bar{A}$. Then there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in A with limit x. Since $A \subseteq B$, $\{x_n\}_{n=1}^{\infty}$ is a sequence in B with limit x; thus $x \in \bar{B}$.

2. Let $x \in \mathring{A}$. Then there exists $r > 0$ such that $B(x,r) \subseteq A$. Since $A \subseteq B$, $B(x,r) \subseteq B$; thus $x \in \mathring{B}$. □

Proposition 3.6. *Let (M,d) be a metric space, and A be a subset of M. Then*

$$x \in \bar{A} \qquad \textit{if and only if} \qquad d(x,A) \equiv \inf\big\{d(x,y)\,\big|\, y \in A\big\} = 0.$$

Proof. "$\Rightarrow$" Suppose that $x \in \bar{A}$. Then there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in A with limit x. By the definition of $d(x,A)$,

$$0 \leqslant d(x,A) = \inf\big\{d(x,y)\,\big|\, y \in A\big\} \leqslant d(x,x_n) \qquad \forall\, n \in \mathbb{N};$$

thus the Sandwich lemma (Lemma 1.34) and Remark 2.43 imply that $d(x,A) = 0$.

"$\Leftarrow$" Suppose $d(x,A) = 0$. By the definition of $d(x,A)$, for all $n \in \mathbb{N}$ there exists $x_n \in A$ such that $d(x,x_n) < d(x,A) + \dfrac{1}{n} = \dfrac{1}{n}$. Therefore, we obtain a sequence $\{x_n\}_{n=1}^{\infty}$ in A such that $\lim\limits_{n\to\infty} x_n = x$; thus $x \in \bar{A}$. □

Remark 3.7. Let (M,d) be a metric space. The function $d(x,A)$ defined in the example above does not satisfy that

$$d(x,y) \leqslant d(x,A) + d(y,A) \qquad \forall\, x,y \in M\,, A \subseteq M\,.$$

However, if $d(A,B) = \inf\big\{d(a,b)\,\big|\, a \in A, b \in B\big\}$, then

$$d(A,B) \leqslant d(x,A) + d(x,B) \qquad \forall\, x \in M\,.$$

The proof of the inequality above is left as an exercise.

Definition 3.8. Let (M,d) be a metric space. A subset A of M is said to be ***dense***（稠密）in another subset B if $A \subseteq B \subseteq \bar{A}$.

Remark 3.9. When A is dense in B, it means that every point in B can be the limit of a sequence in A.

Example 3.10. The rational numbers $\mathbb{Q}$ is dense in the real number system $\mathbb{R}$.

Definition 3.11. Let (M, d) be a metric space, and A be a subset of M. The ***boundary*** of A, denoted by $\text{bd}(A)$ or ∂A, is the intersection of $\bar{A}$ and $\overline{A^\complement}$ ($\partial A = \bar{A} \cap \overline{A^\complement}$).

Remark 3.12. 1. By the definition of limit points of sets, we find that

$$\begin{aligned} x \in \partial A &\Leftrightarrow \exists \{x_n\}_{n=1}^\infty \subseteq A \text{ and } \{y_n\}_{n=1}^\infty \subseteq A^\complement \ni x_n \to x \text{ and } y_n \to x \text{ as } n \to \infty \\ &\Leftrightarrow \forall \varepsilon > 0, B(x, \varepsilon) \cap A \neq \varnothing \text{ and } B(x, \varepsilon) \cap A^\complement \neq \varnothing. \end{aligned}$$

2. $\partial A = \partial(A^\complement)$.

Proposition 3.13. *Let (M, d) be a metric space, and A be a subset of M. Then $\partial A = \bar{A} \backslash \mathring{A}$.*

Proof. If $x \in \partial A$, then $x \in \bar{A} \cap \overline{A^\complement}$; thus for all $\varepsilon > 0$, $B(x, \varepsilon) \cap A^\complement \neq \varnothing$. Therefore, $x \notin \mathring{A}$ which implies that $\partial A \subseteq \bar{A} \backslash \mathring{A}$.

On the other hand, if $x \in \bar{A} \backslash \mathring{A}$, then $x \notin \mathring{A}$; thus for all $\varepsilon > 0$, $B(x, \varepsilon) \nsubseteq A$. As a consequence, for all $\varepsilon > 0$, $B(x, \varepsilon) \cap A^\complement \neq \varnothing$; thus $x \in \overline{A^\complement}$ and this further implies that $x \in \bar{A} \cap \overline{A^\complement} = \partial A$. □

Remark 3.14. 1. If $A \subseteq B$, then in general $\partial A \nsubseteq \partial B$. For example, let $A = \mathbb{Q} \cap [0, 1]$ and $B = [0, 1]$. Then $A \subseteq B$ but $\partial A = [0, 1]$, $\partial B = \{0, 1\}$.

2. It is not always true that $\partial A = \partial(\text{int}(A))$. For example, take $A = [0, 1] \cup \{2\}$. Then $\partial A = \{0, 1, 2\}, \text{int}(A) = (0, 1), \partial(\text{int}(A)) = \{0, 1\}$, so $\partial A \neq \partial(\text{int}(A))$.

3.2 Closed Sets and Open Sets

3.2.1 Closed sets

Definition 3.15. Let (M, d) be a metric space. A subset F of M is said to be ***closed*** (in M) if F contains all its limit points; that is, $F \supseteq \bar{F}$. In other words, F is closed if every convergent sequence $\{x_k\}_{k=1}^\infty \subseteq F$ converges to a limit in F.

Remark 3.16. Let (M, d) be a metric space, and A be a subset of M.

1. By the definition of the closure of sets, $A \subseteq \bar{A}$; thus A is closed if and only if $A = \bar{A}$.

2. By Remark 3.3, A is closed if and only if for all $x \in A^\complement$ there exists $r > 0$ such that $B(x, r) \subseteq A^\complement$.

3. By the definition of the exterior points and 2 above, A is closed if and only if every point of $A^\complement$ is an interior point of $A^\complement$ (or equivalently, $A^\complement = \text{int}(A^\complement)$).

4. If B is a closed set and $A \subseteq B$, then Theorem 3.5 implies that $\bar{A} \subseteq B$.

Example 3.17. Let $a, b \in \mathbb{R}$. The interval $[a, b]$, $(-\infty, a]$, $[b, \infty)$ in $\mathbb{R}$ are closed. This is why $[a, b]$, $(-\infty, a]$ and $[b, \infty)$ are called "closed" intervals in $\mathbb{R}$.

The set $(0, 1] \subseteq \mathbb{R}$ is not closed because it does not contain 0, a limit point of $(0, 1]$.

In general, we have the following

Proposition 3.18. *Let (M, d) be a metric space, $x \in M$ and $r \geqslant 0$.*

1. *The set $B(x, r)^\complement$ is closed.*

2. *The set $\big\{y \in M \,\big|\, d(x, y) \leqslant r\big\}$ is closed.*

Proof. 1. Suppose the contrary that there exists a sequence $\{y_n\}_{n=1}^\infty \subseteq B(x, r)^\complement$ which converges to some $y \in B(x, r)$. Note that $d(x, y) < r$; thus there exists $N > 0$ such that

$$d(y_n, y) < \varepsilon = r - d(x, y) \qquad \text{whenever} \quad n \geqslant N.$$

By the triangle inequality, for $n \geqslant N$ we have

$$d(y_n, x) \leqslant d(y_n, y) + d(y, x) < r - d(x, y) + d(x, y) = r$$

which implies that $y_n \in B(x, r)$ for $n \geqslant N$, a contradiction.

2. Let $A = \big\{y \in M \,\big|\, d(x, y) \leqslant r\big\}$. Suppose the contrary that there exists a sequence $\{y_n\}_{n=1}^\infty \subseteq A$ which converges to some $y \in A^\complement$. Since $d(x, y) > r$, there exists $N > 0$ such that

$$d(y_n, y) < \varepsilon = d(x, y) - r \qquad \text{whenever} \quad n \geqslant N.$$

By the triangle inequality, for $n \geqslant N$ we have

$$d(y_n, x) \geqslant d(x, y) - d(y, y_n) > d(x, y) - \big(d(x, y) - r\big) = r$$

which implies that $y_n \notin A$ for $n \geqslant N$, a contradiction. □

Remark 3.19. When $r = 0$, the set $\big\{y \in M \,\big|\, d(x, y) \leqslant r\big\}$ contains only one point x; thus every set consisting of one single point in M is closed.

Definition 3.20. Let (M,d) be a metric space. For each $x \in M$ and $r > 0$, the set

$$B[x,r] \equiv \{y \in M \,|\, d(x,y) \leqslant r\}$$

is called the closed r-ball about x or the closed ball centered at x with radius r.

Proposition 3.21. *Let (M,d) be a metric space.*

1. *The union of finitely many closed sets is closed.*
2. *The intersection of arbitrary family of closed sets is closed.*
3. *The universal set M and the empty set $\varnothing$ are closed.*

Proof. 1. Let $F_1, \cdots, F_k$ be closed sets in M, $F = \bigcup_{j=1}^{k} F_j$, and $\{x_n\}_{n=1}^{\infty} \subseteq F$ be a convergent sequence with limit $x \in M$. Then there exists $1 \leqslant j_0 \leqslant k$ such that

$$\#\{n \in \mathbb{N} \,|\, x_n \in F_{j_0}\} = \infty;$$

thus $\{n \in \mathbb{N} \,|\, x_n \in F_{j_0}\} = \{n_1, n_2, \cdots, n_k, \cdots\}$, where $n_k < n_{k+1}$ for all $k \in \mathbb{N}$. By Proposition 2.56, the sequence $\{x_{n_k}\}_{k=1}^{\infty}$ converges to x. Since F_{j_0} is closed, $x \in F_{j_0}$; thus $x \in F$. Therefore, every convergent sequence in F converges to a limit in F which shows that F is closed.

2. Let $\mathscr{F} = \{F_\alpha \,|\, F_\alpha \text{ closed in } M,\ \alpha \in I\}$ be a family of closed sets, $F \equiv \bigcap_{\alpha \in I} F_\alpha$, and $\{x_n\}_{n=1}^{\infty} \subseteq F$ be a convergent sequence with limit $x \in M$. Then for each $\alpha \in I$, $\{x_n\}_{n=1}^{\infty}$ in F_α; thus the closedness of F_α implies that $x \in F_\alpha$ for each $\alpha \in I$. Therefore, $x \in F$ which shows that F is closed.

3. Since every consequence in M converges to a limit in M, M must be closed. Since there is no sequence in $\varnothing$, it holds that $\varnothing$ is closed. ▫

Alternative proof of 1 *and* 2. 1. Let $F_1, \cdots, F_k$ be closed sets, $F = \bigcup_{j=1}^{k} F_j$, and $x \in F^{\complement}$. By De Morgan's law,

$$F^{\complement} = M \backslash F = M \backslash \bigcup_{j=1}^{k} F_j = \bigcap_{j=1}^{k} (M \backslash F_j) = \bigcap_{j=1}^{k} F_j^{\complement},$$

so $x \in F_j^{\complement}$ for all $1 \leqslant j \leqslant k$. By Remark 3.16, for each $1 \leqslant j \leqslant k$ there exists $r_j > 0$ such that

$$B(x, r_j) \subseteq F_j^{\complement}.$$

Define $r = \min\{r_1, r_2, \cdots, r_j\}$. Then $r > 0$ and $B(x,r) \subseteq B(x,r_j) \subseteq F_j^{\complement}$ for all $1 \leqslant j \leqslant k$; thus

$$B(x,r) \subseteq \bigcap_{j=1}^{k} F_j^{\complement} \subseteq \big(\bigcup_{j=1}^{k} F_j\big)^{\complement} = F^{\complement}.$$

2. Let $\mathscr{F} = \big\{F_\alpha \,\big|\, F_\alpha \text{ closed in } M,\ \alpha \in I\big\}$ be a family of closed sets, $F \equiv \bigcap_{\alpha\in I} F_\alpha$, and $x \in F$. By De Morgan's law,

$$F^{\complement} = M \backslash \bigcap_{\alpha\in I} F_\alpha = \bigcup_{\alpha\in I} (M\backslash F_\alpha) = \bigcup_{\alpha\in I} F_\alpha^{\complement}$$

so $x \in F_\beta^{\complement}$ for some $\beta \in I$. By Remark 3.16, there exists $r > 0$ such that

$$B(x,r) \subseteq F_\beta^{\complement} \subseteq \bigcup_{\alpha\in I} F_\alpha^{\complement} = \big(\bigcap_{\alpha\in I} F_\alpha\big)^{\complement} = F^{\complement}.$$ ▫

Corollary 3.22. *Every set consisting of finitely many points of a metric space is closed.*

Example 3.23. Let $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, $A \subseteq \mathcal{V}$ be closed, and $B \subseteq \mathcal{V}$ be finite ($\#(B) < \infty$). Then $A + B$ is closed.

Proof. Let $\{\boldsymbol{x}_n\}_{n=1}^{\infty}$ be a convergent sequence in $A+B$ with limit $\boldsymbol{x}$. Then $\boldsymbol{x}_n = \boldsymbol{a}_n + \boldsymbol{b}_n$ for some $\boldsymbol{a}_n \in A$ and $\boldsymbol{b}_n \in B$. Since $\#(B) < \infty$, there exists a point $\boldsymbol{b} \in B$ such that

$$\#\{n \in \mathbb{N} \,|\, \boldsymbol{b}_n = \boldsymbol{b}\} = \infty.$$

Let $\{n \in \mathbb{N} \,|\, \boldsymbol{b}_n = b\} = \{n_1, n_2, \cdots, n_k, \cdots\}$, where $n_k < n_{k+1}$ for all $k \in \mathbb{N}$. Then $\{\boldsymbol{x}_{n_k}\}_{k=1}^{\infty}$ is a subsequence of $\{\boldsymbol{x}_n\}$. By Proposition 2.56, $\{\boldsymbol{x}_{n_k}\}_{k=1}^{\infty}$ converges to $\boldsymbol{x}$; thus the fact that $\boldsymbol{a}_{n_k} = \boldsymbol{x}_{n_k} - \boldsymbol{b}$ shows that $\{\boldsymbol{a}_{n_k}\}_{k=1}^{\infty}$ converges to a limit $\boldsymbol{a}$ and $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{b}$. The closedness of A further implies that $\boldsymbol{a} \in A$; thus $\boldsymbol{x} \in A + B$.

In fact, $A + B = \bigcup_{\boldsymbol{b}\in B} (\boldsymbol{b} + A)$. It should be clear that $\boldsymbol{b} + A$ is closed if A is closed; thus we conclude that $A + B$ is open by Proposition 3.21. ▫

Theorem 3.24. *Let (M,d) be a metric space, and A be a subset of M. Then $\bar{A}$ is closed.*

Proof. Let $x \in \bar{A}^{\complement}$ be given. Remark 3.3 implies that there exists $\varepsilon > 0$ such that $B(x,\varepsilon) \cap A = \varnothing$ or equivalently, $A \subseteq B(x,\varepsilon)^{\complement}$. By Proposition 3.18, $B(x,\varepsilon)^{\complement}$ is closed; thus 4 of Remark 3.16 implies that $\bar{A} \subseteq B(x,\varepsilon)^{\complement}$. Therefore, $B(x,\varepsilon) \subseteq \bar{A}^{\complement}$; thus we established that every point in $\bar{A}^{\complement}$ is an interior point of $\bar{A}^{\complement}$. 4 of Remark 3.16 then shows that $\bar{A}$ is closed. ▫

Remark 3.25. Let (M, d) be a metric space, and A be a subset of M.

1. By the definition of the boundary of sets, Proposition 3.21 and Theorem 3.24 imply that ∂A is closed.

2. By 4 of Remark 3.16, every closed set containing A contains $\bar{A}$; thus the closure of A is the smallest closed set containing A; that is, $\bar{A} = \bigcap\limits_{\substack{A \subseteq F \\ F \text{ closed}}} F$.

Definition 3.26. Let (M, d) be a metric space. A subset A of M is said to be complete if the metric space (A, d) is complete. In other words, A is complete if every Cauchy sequence in A converges to a limit in A.

Theorem 3.27. *Let (M, d) be a complete metric space, and A be a subset of M. Then A is complete if and only if A is closed in M.*

Proof. "$\Rightarrow$" Let $\{x_k\}_{k=1}^{\infty}$ be a convergent sequence in A with limit x. Then Proposition 2.58 implies that $\{x_k\}_{k=1}^{\infty}$ is a Cauchy sequence in (M, d). Since $\{x_k\}_{k=1}^{\infty} \subseteq A$, we find that $\{x_k\}_{k=1}^{\infty}$ is a Cauchy sequence in (A, d); thus the completeness of (A, d) implies that $\{x_k\}_{k=1}^{\infty}$ converges to $y \in A$. By Proposition 2.42, the limit is unique; thus $x = y$ which implies that $x \in A$. Therefore, A is closed.

"$\Leftarrow$" Let $\{x_k\}_{k=1}^{\infty}$ be a Cauchy sequence in A. Then

$$\forall \varepsilon > 0, \exists N > 0 \ni d(x_n, x_m) < \varepsilon \quad \text{whenever} \quad n, m \geqslant N .$$

Therefore, $\{x_k\}_{k=1}^{\infty}$ is a Cauchy in (M, d). Since (M, d) is complete, there exists $x \in M$ such that $x_k \to x$ as $k \to \infty$. By the closedness of A, we must have $x \in A$; thus every Cauchy sequence in A converges to a limit in A. □

3.2.2 Open sets

Definition 3.28. Let (M, d) be a metric space. A set $U \subseteq M$ is said to be ***open*** (in M) if every point in U is an interior point of U; that is, $U \subseteq \mathring{U}$. In other words,

$$U \text{ is open (in } M) \quad \Leftrightarrow \quad \forall\, x \in U,\ \exists\, r > 0 \ni B(x, r) \subseteq U .$$

Remark 3.29. Let (M, d) be a metric space, and A be a subset of M.

1. By the definition of the interior of sets, $\mathring{A} \subseteq A$; thus A is open if and only if $A = \mathring{A}$.

2. By 3 of Remark 3.16, A is open if and only if $A^\complement$ is closed and A is closed if and only if $A^\complement$ is open.

3. The statement above does **NOT** implies that a set in a metric space is either open or closed. There are still sets which is neither open nor closed.

4. If B is an open set and $B \subseteq A$, then Theorem 3.5 implies that $B \subseteq \mathring{A}$.

Example 3.30. The interval (a,b) in $\mathbb{R}$ is open since $\text{int}((a,b)) = (a,b)$. This is why (a,b) is called an open interval in Calculus.

Example 3.31. The set $A = \{(a,b) \in \mathbb{R}^2 \,|\, 0 < a < 1\}$ is open: given $x = (a,b) \in A$, take $r = \min\{1-a, a\}$, then $B(x,r) \subseteq A$.

On the other hand, the set $A = \{(a,b) \in \mathbb{R}^2 \,|\, 0 < a \leqslant 1\}$ is not open: let $x = (1,0)$, then for each $r > 0$, $B(x,r) \nsubseteq A$ since the point $\big(1+\frac{r}{2}, 0\big) \in B(x,r)$ but $\big(1+\frac{r}{2}, 0\big) \notin A$.

The following proposition is a direct consequence of Proposition 3.18 and Remark 3.29.

Proposition 3.32. *Every r-ball in a metric space is open.*

Alternative proof. Let (M,d) be a metric space, and $B(x,r)$ be an r-ball in M. We would like to show that for each $y \in B(x,r)$, there exists $\delta > 0$ such that $B(y,\delta) \subseteq B(x,r)$. Let $\delta = r - d(x,y)$. Then $\delta > 0$ and if $z \in B(y,\delta)$, we have

$$d(z,x) \leqslant d(z,y) + d(y,x) < \delta + d(y,x) = r;$$

thus $z \in B(x,r)$. □

Proposition 3.33. *Let (M,d) be a metric space.*

1. *The intersection of finitely many open sets is open.*

2. *The union of arbitrary family of open sets is open.*

3. *The empty set $\varnothing$ and the universal set M are open.*

Proof. 1. Let $U_1, \cdots, U_k$ be open sets, and $U = \bigcap\limits_{j=1}^{k} U_j$. Then by De Morgan's law,

$$U^\complement = M\backslash U = M\backslash \bigcap_{j=1}^{k} U_j = \bigcup_{j=1}^{k} (M\backslash U_j) = \bigcup_{j=1}^{k} {U_j}^\complement.$$

Since U_j is open, ${U_j}^\complement$ is closed. By Proposition 3.21, $\bigcup\limits_{j=1}^{k} {U_j}^\complement$ is closed.

2. Let $\mathscr{F} = \big\{U_\alpha \,\big|\, U_\alpha \text{ open in } M,\ \alpha \in I\big\}$ be a family of open sets, and $U \equiv \bigcup\limits_{\alpha\in I} U_\alpha$. Then by De Morgan's law,
$$U^{\complement} = M\backslash\bigcup_{\alpha\in I}U_\alpha = \bigcap_{\alpha\in I}(M\backslash U_\alpha) = \bigcup_{\alpha\in I}U_\alpha^{\complement}$$
which implies that $U^{\complement}$ is the intersection of a family of closed sets $\big\{U_\alpha^{\complement}\big\}_{\alpha\in I}$. By Proposition 3.21 we conclude that $U^{\complement}$ is closed or equivalently, U is open. ▫

Alternative proof of 1 *and* 2.

1. Let $U_1, U_2, \cdots, U_k$ be open sets in M, and $U \equiv \bigcap\limits_{i=1}^{k} U_i$. If $y \in U$, then $y \in U_i$ for all $1 \leqslant i \leqslant k$. Since U_i is open, there exist $\delta_i > 0$ such that $B(y, \delta_i) \subseteq U_i$. Let $\delta = \min\{\delta_1, \cdots, \delta_k\}$. Next we show that $B(y, \delta) \subseteq U$ to conclude that U is open.

   Let $z \in B(y, \delta)$. Then $d(y, z) < \delta \leqslant \delta_i$ for all $1 \leqslant i \leqslant k$. Therefore, $z \in B(y, \delta_i)$ for all $1 \leqslant i \leqslant k$ which shows that $z \in U_i$ for all $1 \leqslant i \leqslant k$; thus $z \in \bigcap\limits_{i=1}^{k} U_i \equiv U$.

2. Let $\mathscr{F} = \big\{U_\alpha \,\big|\, U_\alpha \text{ open in } M,\ \alpha \in I\big\}$ be a family of open sets, and $U \equiv \bigcup\limits_{\alpha\in I} U_\alpha$. If $y \in U$, then $y \in U_\beta$ for some $\beta \in I$. Since U_β is open, there exists $\delta > 0$ such that $B(y, \delta) \subseteq U_\beta$; thus $B(y, \delta) \subseteq \bigcup\limits_{\alpha\in I} U_\alpha \equiv U$. ▫

Remark 3.34. Infinite intersection of open sets need not be open:

1. Take $A_k = \big(-\dfrac{1}{k}, \dfrac{1}{k}\big)$, then $\bigcap\limits_{k=1}^{\infty} A_k = \{0\}$ which is not open.

2. Let $U_k = (-2 - \dfrac{1}{k}, 2 + \dfrac{1}{k}) \subseteq \mathbb{R}$. Then $A = \bigcap\limits_{k=1}^{\infty} U_k \supseteq [-2, 2]$. Moreover, if $x \notin [-2, 2]$, then $\exists\, k \in \mathbb{N} \ni x \notin U_k$ $\Big($If $x > 2$, $\dfrac{1}{k} < \dfrac{x-2}{2}$. If $x < -2$, $\dfrac{1}{k} < \dfrac{-x-2}{2}\Big)$. Therefore, $\bigcap\limits_{k=1}^{\infty} U_k = [-2, 2]$.

Example 3.35. Let $(\mathcal{V}, \|\cdot\|)$ be a normed vector space. If A, B are subsets of $\mathcal{V}$ and A is open, the set $A + B$ is open.

Proof. Let $\boldsymbol{y} \in A + B$. Then $\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{b}$ for some $\boldsymbol{a} \in A$, $\boldsymbol{b} \in B$. Since A is open, there exists $\delta > 0$ such that $B(\boldsymbol{a}, \delta) \subseteq A$.

We next show that $B(\boldsymbol{y}, \delta) \subseteq A + B$. Let $\boldsymbol{z} \in B(\boldsymbol{y}, \delta)$. Then $\|\boldsymbol{z} - \boldsymbol{y}\| < \delta$. Since $\boldsymbol{z} = \boldsymbol{b} + (\boldsymbol{z} - \boldsymbol{b})$, if we can show that $\boldsymbol{z} - \boldsymbol{b} \in A$, then $\boldsymbol{z} \in A + B$. Nevertheless, we have

$$\|(\boldsymbol{z} - \boldsymbol{b}) - \boldsymbol{a}\| = \|\boldsymbol{z} - \boldsymbol{a} - \boldsymbol{b}\| = \|\boldsymbol{z} - \boldsymbol{y}\| < \delta$$

which implies that $\boldsymbol{z} - \boldsymbol{b} \in B(\boldsymbol{a}, \delta) \subseteq A$. □

Since $A + B = \bigcup_{\boldsymbol{b} \in B} (\boldsymbol{b} + A)$ and it should be clear that $\boldsymbol{b} + A$ is open if A is open, we conclude by Proposition 3.33 that $A + B$ is open.

Theorem 3.36. *Let (M, d) be a metric space, and A be a subset of M. Then $\mathring{A}$ is open.*

Proof. For each $x \in \mathring{A}$, let $\varepsilon_x > 0$ denote a number such that $B(x, \varepsilon_x) \subseteq A$. We would like to show that $\mathring{A} = \bigcup_{x \in \mathring{A}} B(x, \varepsilon_x)$, and the theorem is then a direct consequence of Proposition 3.33.

1. "$\subseteq$": trivial.
2. "$\supseteq$": Let $y \in \bigcup_{x \in \mathring{A}} B(x, \varepsilon_x)$. By the definition of the union of family of sets, there exists $x \in \mathring{A}$ such that $y \in B(x, \varepsilon_x)$. Let $\delta = \varepsilon_x - d(x, y)$. Then $\delta > 0$ and if $z \in B(y, \delta)$,

$$d(z, x) \leqslant d(z, y) + d(y, x) < \delta + d(y, x) = \varepsilon_x$$

which implies that $B(y, \delta) \subseteq B(x, \varepsilon_x) \subseteq A$. Therefore, $y \in \mathring{A}$. □

Remark 3.37. Let (M, d) be a metric space, and A be a subset of M. By 4 of Remark 3.29, every open set contained inside A is contained inside $\mathring{A}$; thus the interior of A is the largest open set contained inside A; that is, $\mathring{A} = \bigcup_{\substack{A \supseteq U \\ U \text{ open}}} U$.

Remark 3.38. In a metric space (M, d), it is **NOT** always true that $\text{int}\big(B[x, R]\big) = B(x, R)$ or $\text{cl}(B(x, R)) = B[x, R]$. For example, we consider the discrete metric

$$d_0(x, y) = \begin{cases} 1 & \text{if } x \neq y, \\ 0 & \text{if } x = 0. \end{cases}$$

Let $R = 1$, and fix $x \in M \neq \varnothing$. Then $B[x, 1] = M$ and $B(x, 1) = \{x\}$. Since every set in (M, d_0) is both closed and open, we find that

$$\text{int}(B[x, 1]) = M \quad \text{and} \quad \text{cl}(B(x, 1)) = \{x\};$$

thus as long as M has more than one point, we have $\text{int}\big(B[x, 1]\big) \neq B(x, 1)$ and $\text{cl}\big(B(x, 1)\big) \neq B[x, 1]$. We also note that in (M, d_0) the boundary of every set is empty.

3.3 Compactness（緊緻性）

In this section, we investigate a property similar to the Bolzaon-Weierstrass Property.

Definition 3.39. Let (M, d) be a metric space. A subset $K \subseteq M$ is called ***sequentially compact*** if every sequence in K has a subsequence that converges to a point in K.

Definition 3.40. Let (M, d) be a metric space. A subset A of M is said to be bounded if A is contained in some r-ball. In other words, A is bounded if there exists $x \in M$ and $r > 0$ such that $A \subseteq B(x, r)$.

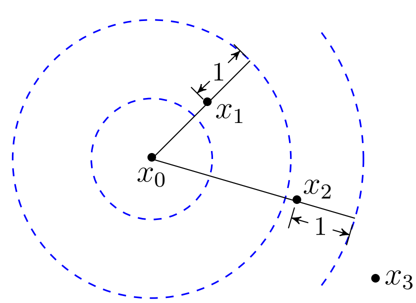
Example 3.41. Each closed and bounded set in $(\mathbb{R}^n, \|\cdot\|_2)$ is sequentially compact. This is a direct consequence of the Bolzano-Weierstrass Theorem (Theorem 2.51) and the definition of the closedness of sets.

Theorem 3.42. *Let (M, d) be a metric space, and $K \subseteq M$ be sequentially compact. Then K is closed and bounded.*

Proof. For closedness, assume that $\{x_k\}_{k=1}^{\infty} \subseteq K$ and $x_k \to x$ as $k \to \infty$. By the definition of sequential compactness, there exists $\{x_{k_j}\}_{j=1}^{\infty}$ converging to a point $y \in K$. By Proposition 2.56, $x = y$; thus $x \in K$.

For boundedness, assume the contrary that for all $x_0 \in M$ and $R > 0$, there exists $y \in K$ such that $d(x_0, y) \geqslant R$. Fix $x_0 \in M$. There exists $x_1 \in K$ such that $d(x_0, x_1) \geqslant 1$. Having x_1, there exists $x_2 \in K$ such that $d(x_2, x_0) \geqslant 1 + d(x_1, x_0)$. Continuing this process, we obtain a sequence $\{x_k\}_{k=1}^{\infty}$ in K such that

$$d(x_k, x_0) \geqslant 1 + d(x_{k-1}, x_0) \quad \forall\, k \in \mathbb{N}.$$



Then any subsequence of $\{x_k\}_{k=1}^{\infty}$ cannot be Cauchy since $d(x_k, x_\ell) \geqslant |k - \ell|$ for all $k, \ell \in \mathbb{N}$; thus $\{x_k\}_{k=1}^{\infty}$ has no convergent subsequence, a contradiction. ∎

Remark 3.43. Example 3.41 and Theorem 3.42 together imply that in $(\mathbb{R}^n, \|\cdot\|_2)$,

$$\text{sequentially compact} \quad \Leftrightarrow \quad \text{closed and bounded}.$$

This result is called the ***Heine-Borel*** Theorem.

In fact, if $\mathcal{V}$ is a **finite dimensional** vector space over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\} \subseteq \mathcal{V}$ is a basis for $\mathcal{V}$; that is, every $\boldsymbol{x} \in \mathcal{V}$ can be uniquely expressed as

$$\boldsymbol{x} = x^{(1)}\mathbf{e}_1 + x^{(2)}\mathbf{e}_2 + \cdots + x^{(n)}\mathbf{e}_n, \qquad x^{(k)} \in \mathbb{F} \text{ for } 1 \leqslant k \leqslant n.$$

Define $\|\boldsymbol{x}\|_2 = \Big(\sum\limits_{i=1}^n |x^{(i)}|^2\Big)^{\frac{1}{2}}$ (which is a norm by Example 2.28). Then a subset K of $\mathcal{V}$ is sequentially compact in $(\mathcal{V}, \|\cdot\|_2)$ if and only if K is closed and bounded. Note that by Theorem 3.42 it suffices to show the "if" direction. Let $\{\boldsymbol{x}_k\}_{k=1}^\infty$ be a sequence in K. Write $\boldsymbol{x}_k = \sum\limits_{i=1}^n x_k^{(i)}\mathbf{e}_i$, and define $\boldsymbol{v}_k = \big(x_k^{(1)}, x_k^{(2)}, \cdots, x_k^{(n)}\big)$. Then $\{\boldsymbol{v}_k\}_{k=1}^\infty$ is a sequence in $\mathbb{F}^n$. Since $\{\boldsymbol{x}_k\}_{k=1}^\infty$ is bounded, there exists $M > 0$ such that

$$\|\boldsymbol{x}_k\|_2 \leqslant M \qquad \forall\, k \in \mathbb{N};$$

thus $\|\boldsymbol{v}_k\|_2 \leqslant M$ (here $\|\boldsymbol{v}_k\|_2$ is the usual norm of $\boldsymbol{v}_k$ on $\mathbb{F}^n$) for all $k \in \mathbb{N}$. By the Bolzano-Weierstrass Theorem (Theorem 2.51 and Remark 2.52), there exists a subsequence $\{\boldsymbol{v}_{k_j}\}_{j=1}^\infty$ such that $\{\boldsymbol{v}_{k_j}\}_{j=1}^\infty$ converges to some $\boldsymbol{v} \in \mathbb{F}^n$. Let $\boldsymbol{v} = \big(x^{(1)}, x^{(2)}, \cdots, x^{(n)}\big)$ and $\boldsymbol{x} = x^{(1)}\mathbf{e}_1 + \cdots + x^{(n)}\mathbf{e}_n$. Then

$$\big\|\boldsymbol{x}_{k_j} - \boldsymbol{x}\big\|_2 = \Big(\sum_{i=1}^n \big|x_{k_j}^{(i)} - x^{(i)}\big|^2\Big)^{\frac{1}{2}} = \|\boldsymbol{v}_{k_j} - \boldsymbol{v}\|_2 \to 0 \;\text{ as }\; j \to \infty,$$

and the closedness of K implies that $\boldsymbol{x} \in K$, so we establish that K is sequentially compact in $(\mathcal{V}, \|\cdot\|_2)$ if K is closed and bounded.

Example 3.44. Let $A = [0,1] \cup (2,3] \subseteq (\mathbb{R}, |\cdot|)$. Since A is not closed, A is not sequentially compact.

Corollary 3.45. *If $K \subseteq \mathbb{R}$ is sequentially compact, then* $\inf K \in K$ *and* $\sup K \in K$.

Proof. By Theorem 3.42, K must be closed and bounded. Therefore, $\inf K \in \mathbb{R}$. Then for each $n \in \mathbb{N}$, there exists $x_n \in K$ such that $\inf K \leqslant x_n < \inf K + \dfrac{1}{n}$. Since $\{x_n\}_{n=1}^\infty$ is a bounded sequence in $\mathbb{R}$, the Bolzano-Weierstrass property of $\mathbb{R}$ implies that there is a subsequence $\{x_{n_k}\}_{k=1}^\infty$ and $x \in \mathbb{R}$ such that $\lim\limits_{k\to\infty} x_{n_k} = x$. Note that $x = \inf K$, and by the closedness of K, $x \in K$. The proof of $\sup K \in K$ is similar. ◻

Definition 3.46. Let (M,d) be a metric space. A subset $A \subseteq M$ is called ***totally bounded*** if for each $r > 0$, there exists $\{x_1, \cdots, x_N\} \subseteq M$ such that

$$A \subseteq \bigcup_{i=1}^{N} B(x_i, r).$$

Remark 3.47. In a general metric space (M,d), a bounded set might not be totally bounded. For example, consider the metric space (M,d) with the discrete metric, and $A \subseteq M$ be a set having infinitely many points. Then A is bounded since $A \subseteq B(x,2)$ for any $x \in M$; however, A is not totally bounded since A cannot be covered by finitely many balls with radius $\frac{1}{2}$.

Proposition 3.48. *Let (M,d) be a metric space, and $A \subseteq M$ be totally bounded. Then A is bounded. In other words, totally bounded sets are bounded.*

Proof. By total boundedness, there exists $\{y_1, \cdots, y_N\} \subseteq M$ such that $A \subseteq \bigcup_{i=1}^{N} B(y_i, 1)$. Let $x_0 = y_1$ and $R = \max\big\{d(x_0, y_2), \cdots, d(x_0, y_N)\big\} + 1$. Then if $z \in A$, $z \in B(y_j, 1)$ for some $j = 1, \cdots, N$, and

$$d(z, x_0) \leqslant d(z, y_j) + d(y_j, x_0) < 1 + d(x_0, y_j) \leqslant R$$

which implies that $A \subseteq B(x_0, R)$. Therefore, A is bounded. □

Example 3.49. Every bounded set in $(\mathbb{R}^n, \|\cdot\|_2)$ is totally bounded **(Check!)**. In particular, the set $\{1\} \times [1,2]$ in $(\mathbb{R}^2, \|\cdot\|_2)$ is totally bounded.

On the other hand, let $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$d(x,y) = \begin{cases} |x_1 - y_1| & \text{if } x_2 = y_2\,, \\ |x_1 - y_1| + |x_2 - y_2| + 1 & \text{if } x_2 \neq y_2\,, \end{cases} \quad \text{where } x = (x_1, x_2) \text{ and } y = (y_1, y_2).$$

Then $(\mathbb{R}^2, d)$ is also a metric space (exercise). The set $\{1\} \times [1,2]$ is bounded **(Check!)** but not totally bounded. In fact, consider open ball with radius $\frac{1}{2}$:

$$\begin{aligned} y \in B\big(x, \tfrac{1}{2}\big) &\Leftrightarrow \|x - y\| < \frac{1}{2} \Leftrightarrow |x_1 - y_1| < \frac{1}{2} \text{ and } x_2 = y_2 \\ &\Leftrightarrow y_1 \in \big(x_1 - \tfrac{1}{2}, x_1 + \tfrac{1}{2}\big) \text{ and } x_2 = y_2\,. \end{aligned}$$

In other words,

$$B\big(x, \tfrac{1}{2}\big) = \big(x_1 - \tfrac{1}{2}, x_1 + \tfrac{1}{2}\big) \times \{x_2\};$$

thus one cannot cover $\{1\} \times [1,2]$ by the union of finitely many balls with radius $\frac{1}{2}$.

Example 3.50. Let ℓ^∞ denote the collection of all bounded real sequences (cf. Example 2.20); that is,

$$\ell^\infty = \big\{\{x_k\}_{k=1}^\infty \subseteq \mathbb{R} \,\big|\, \text{for some } M > 0, |x_k| \leqslant M \text{ for all } k\big\}.$$

The number $\sup\limits_{k\geqslant 1} |x_k| \equiv \sup\{|x_1|, |x_2|, \cdots, |x_k|, \cdots\} < \infty$ is denoted by $\big\|\{x_k\}_{k=1}^\infty\big\|_\infty$ $\Big($for example, if $x_k = \dfrac{(-1)^k}{k}$, then $\big\|\{x_k\}_{k=1}^\infty\big\|_\infty = 1\Big)$. Then $(\ell^\infty, \|\cdot\|_\infty)$ is a Banach space (left as an exercise). Define

$$\begin{aligned}
A &= \Big\{\{x_k\}_{k=1}^\infty \in \ell^\infty \,\Big|\, |x_k| \leqslant \frac{1}{k}\Big\},\\
B &= \big\{\{x_k\}_{k=1}^\infty \in \ell^\infty \,\big|\, x_k \to 0 \text{ as } k \to \infty\big\},\\
C &= \big\{\{x_k\}_{k=1}^\infty \in \ell^\infty \,\big|\, \text{the sequence } \{x_k\}_{k=1}^\infty \text{ converges}\big\},\\
D &= \Big\{\{x_k\}_{k=1}^\infty \in \ell^\infty \,\Big|\, \sup_{k\geqslant 1} |x_k| = 1\Big\} \quad (\text{the unit sphere in } (\ell^\infty, \|\cdot\|)).
\end{aligned}$$

The closedness of A (which implies the completeness of $(A, \|\cdot\|_\infty)$) is left as an exercise. We show that A is totally bounded.

Let $r > 0$ be given. Then there exists $N > 0$ such that $\dfrac{1}{N} < r$. Define

$$\begin{aligned}
E = \Big\{\{x_k\}_{k=1}^\infty \,\Big|\, & x_1 = \frac{i_1}{N+1}, x_2 = \frac{i_2}{N+1}, \cdots, x_{N-1} = \frac{i_{N-1}}{N+1} \text{ for some}\\
& i_1, \cdots, i_{N-1} = -N, -N+1, \cdots, N-1, N, \text{ and } x_k = 0 \text{ if } k \geqslant N+1\Big\}.
\end{aligned}$$

Then

1. $\#E < \infty$. In fact, $\#E = (2N+1)^{N-1} < \infty$.
2. $A \subseteq \displaystyle\bigcup_{\{x_k\}_{k=1}^\infty \in E} B\big(\{x_k\}_{k=1}^\infty, \frac{1}{N}\big) \subseteq \bigcup_{\{x_k\}_{k=1}^\infty \in E} B\big(\{x_k\}_{k=1}^\infty, r\big)$.

Therefore, A is totally bounded.

On the other hand, B and C are not bounded; thus not totally bounded by Proposition 3.48. D is bounded but not totally bounded. In fact, D cannot be covered by the union of finitely many balls with radius $\dfrac{1}{2}$ since each ball with radius $\dfrac{1}{2}$ contains at most one of the points from the subset $\Big\{\{x_j^{(k)}\}_{j=1}^\infty\Big\}_{k=1}^\infty \subseteq D$, where for each k

$$\{x_j^{(k)}\}_{j=1}^\infty = \{\underbrace{0, \cdots, 0,}_{(k-1)\text{ terms}} 1, 0, \cdots\};$$

that is, $x_j^{(k)} = \delta_{kj}$, the kronecker delta.

Proposition 3.51. *Let (M,d) be a metric space, and $T \subseteq M$ be totally bounded. If $S \subseteq T$, then S is totally bounded. In other words, subsets of totally bounded sets are totally bounded.*

Proof. Let $r > 0$ be given. By the total boundedness of T, there exists $\{x_1, \cdots, x_N\} \subseteq M$ such that
$$S \subseteq T \subseteq \bigcup_{i=1}^{N} B(x_i, r). \qquad \square$$

Proposition 3.52. *Let (M,d) be a metric space, and $A \subseteq M$. Then A is totally bounded if and only if for all $r > 0$, there exists $\{y_1, \cdots, y_N\} \subseteq A$ such that $A \subseteq \bigcup_{i=1}^{N} B(y_i, r)$.*

Proof. It suffices to show the "only if" part. Let $r > 0$ be given. Since A is totally bounded,
$$\exists \{y_1, \cdots, y_N\} \subseteq M \ni A \subseteq \bigcup_{i=1}^{N} B\big(y_i, \frac{r}{2}\big).$$
W.L.O.G., we may assume that for each $i = 1, \cdots, N$, $B\big(y_i, \frac{r}{2}\big) \cap A \neq \varnothing$. Then for each $i = 1, \cdots, N$, there exists $x_i \in B\big(y_i, \frac{r}{2}\big) \cap A$ which implies that
$$A \subseteq \bigcup_{i=1}^{N} B\big(y_i, \frac{r}{2}\big) \subseteq \bigcup_{i=1}^{N} B(x_i, r)$$
since $B\big(y_i, \frac{r}{2}\big) \subseteq B(x_i, r)$ for all $i = 1, \cdots, N$. $\square$

Theorem 3.53. *Let (M,d) be a metric space, and K be a subset of M. Then K is sequentially compact if and only if K is totally bounded and complete.*

Proof. "$\Rightarrow$" Assume that K is sequentially compact. For the total boundedness, suppose the contrary that there is an $r > 0$ such that any finite set $\{y_1, \cdots, y_n\} \subseteq K$, $K \nsubseteq \bigcup_{i=1}^{n} B(y_i, r)$. This implies that we can choose a sequence $\{x_k\}_{k=1}^{\infty} \subseteq K$ such that
$$x_{k+1} \in K \setminus \bigcup_{i=1}^{k} B(x_i, r).$$
Then $\{x_k\}_{k=1}^{\infty}$ is a sequence in K without convergent subsequence since $d(x_k, x_\ell) > r$ for all $k, \ell \in \mathbb{N}$ and $k \neq \ell$.

Next we show that K is complete. Let $\{x_k\}_{k=1}^{\infty} \subseteq K$ be a Cauchy sequence. By sequential compactness of K, there is a subsequence $\{x_{k_j}\}_{j=1}^{\infty}$ converging to a point $x \in K$. By Proposition 2.58, $\{x_k\}_{k=1}^{\infty}$ also converges to x; thus every Cauchy sequence in K converges to a point in K.

"$\Leftarrow$" The proof of this direction is similar to the proof of Theorem 1.79 and we proceed as follows. Let $\{x_k\}_{k=1}^{\infty}$ be a sequence in $T_0 \equiv K$. Since K is totally bound, there exists $\big\{y_1^{(1)}, \cdots y_{N_1}^{(1)}\big\} \subseteq K$ such that
$$T_0 \equiv K \subseteq \bigcup_{i=1}^{N_1} B(y_i^{(1)}, 1)\,.$$
One of these $B(y_i^{(1)}, 1)$'s must contain infinitely many x_k's; that is, there exists $1 \leqslant \ell_1 \leqslant N_1$ such that $\#\big\{k \in \mathbb{N} \,\big|\, x_k \in B(y_{\ell_1}^{(1)}, 1)\big\} = \infty$. Define $T_1 = K \cap B(y_{\ell_1}^{(1)}, 1)$. Then T_1 is also totally bounded by Proposition 3.51, so there exists $\big\{y_1^{(2)}, \cdots, y_{N_2}^{(2)}\big\} \subseteq T_1$ such that
$$T_1 \subseteq \bigcup_{i=1}^{N_2} B\big(y_i^{(2)}, \frac{1}{2}\big)\,.$$
Suppose that $\#\big\{k \in \mathbb{N} \,\big|\, x_k \in B\big(y_{\ell_2}^{(2)}, \frac{1}{2}\big)\big\} = \infty$ for some $1 \leqslant \ell_2 \leqslant N_2$. Define $T_2 = T_1 \cap B\big(y_{\ell_2}^{(2)}, \frac{1}{2}\big)$. We continue this process, and obtain that for all $n \in \mathbb{N}$,

(1) there exists $\big\{y_1^{(n)}, \cdots, y_{N_n}^{(n)}\big\} \subseteq T_{n-1}$ such that
$$T_{n-1} \subseteq \bigcup_{i=1}^{N_n} B\big(y_i^{(n)}, \frac{1}{n}\big)\,.$$

(2) $T_n = T_{n-1} \cap B(y_{\ell_n}^{(n)}, \frac{1}{n})$, where $1 \leqslant \ell_n \leqslant N_n$ is chosen so that
$$\#\Big\{k \in \mathbb{N} \,\Big|\, x_k \in B\big(y_{\ell_n}^{(n)}, \frac{1}{n}\big)\Big\} = \infty\,. \tag{3.3.1}$$

Pick an $k_1 \in \big\{k \in \mathbb{N} \,\big|\, x_k \in B\big(y_{\ell_1}^{(1)}, 1)\big\}$, and $k_j \in \Big\{k \in \mathbb{N} \,\Big|\, x_k \in B\big(y_{\ell_j}^{(j)}, \frac{1}{j}\big)\Big\}$ such that $k_{j+1} > k_j$ for all $j \in \mathbb{N}$. We note such k_j always exists because of (3.3.1). Then $\big\{x_{k_j}\big\}_{j=1}^{\infty}$ is a subsequence of $\{x_k\}_{k=1}^{\infty}$, and $x_{k_j} \in T_j \subseteq K$ for all $j \in \mathbb{N}$.

Claim: $\big\{x_{k_j}\big\}_{j=1}^{\infty}$ is a Cauchy sequence.

Proof of claim: Let $\varepsilon > 0$ be given, and $N > 0$ be large enough so that $\frac{1}{N} < \frac{\varepsilon}{2}$. Then if $j \geqslant N$, we must have $x_{k_j} \in B\big(y_{\ell_N}^{(N)}, \frac{1}{N}\big)$; thus we conclude that if $n, m \geqslant N$, the triangle inequality implies that
$$d\big(x_{k_n}, x_{k_m}\big) \leqslant d\big(x_{k_n}, y_{\ell_N}^{(N)}\big) + d\big(x_{k_m}, y_{\ell_N}^{(N)}\big) < \frac{1}{N} + \frac{1}{N} < \varepsilon.$$
Since (K, d) is complete, the Cauchy sequence $\big\{x_{k_j}\big\}_{j=1}^{\infty}$ converges to a point in K. $\square$

Finally we introduce the concept of compact sets in a metric space in the remaining part of the section. First we need to talk about open cover of sets.

Definition 3.54. Let (M, d) be a metric space, and A be a subset of M. A ***cover*** of A is a collection of sets $\{U_\alpha\}_{\alpha\in I}$ satisfying that $A \subseteq \bigcup_{\alpha\in I} U_\alpha$. It is an ***open cover*** of A if U_α is open for all $\alpha \in I$. A ***subcover*** of a given cover $\{U_\alpha\}_{\alpha\in I}$ is a sub-collection $\{U_\alpha\}_{\alpha\in J}$, where $J \subseteq I$, satisfying that $A \subseteq \bigcup_{\alpha\in I} U_\alpha$. It is a ***finite subcover*** if $\#J < \infty$.

Example 3.55. The collection $\big\{(-k, k) \,\big|\, k \in \mathbb{N}\big\}$ is an open cover of $\mathbb{R}$, and $\big\{(-2k, 2k) \,\big|\, k \in \mathbb{N}\big\}$ is a subcover of $\big\{(-k, k) \,\big|\, k \in \mathbb{N}\big\}$.

The so-called compact sets is defined in the following

Definition 3.56. Let (M, d) be a metric space. A subset $K \subseteq M$ is called ***compact*** if every open cover of K possesses a finite subcover; that is, $K \subseteq M$ is compact if

$$\big(\forall \text{ open cover } \{U_\alpha\}_{\alpha\in I} \text{ of } K\big)\big(\exists\, J \subseteq I \wedge \#J < \infty\big)\Big(K \subseteq \bigcup_{\alpha\in J} U_\alpha\Big).$$

Example 3.57. Let $A = \{0\} \cup \Big\{1, \dfrac{1}{2}, \cdots, \dfrac{1}{n}, \cdots\Big\}$, and $\{U_\alpha\}_{\alpha\in I}$ be a given open cover of A. Then $0 \in U_\beta$ for all $\beta \in I$. By the fact that U_β is open, there exists $\varepsilon > 0$ such that $B(0, \varepsilon) \subseteq U_\beta$. Let $N \in \mathbb{N}$ satisfy $\dfrac{1}{N} < \varepsilon$. Then

$$\Big\{\frac{1}{N}, \frac{1}{N+1}, \cdots\Big\} = \Big\{\frac{1}{k} \,\Big|\, k \geqslant N\Big\} \subseteq U_\beta.$$

On the other hand, for each $1 \leqslant k \leqslant N-1$ there exists $\alpha_k \in I$ such that $\dfrac{1}{k} \in U_{\alpha_k}$. Therefore,

$$A = \Big(\{0\} \cup \Big\{\frac{1}{N}, \frac{1}{N+1}, \cdots\Big\}\Big) \cup \Big\{1, \frac{1}{2}, \cdots, \frac{1}{N-1}\Big\} \subseteq U_\beta \cup \bigcup_{k=1}^{N-1} U_{\alpha_k};$$

thus we obtain a finite subcover for a given open cover. Therefore, A is compact.

In general, if $\{x_k\}_{k=1}^\infty$ is a convergent sequence in a metric space with limit x, then the set $A = \{x_1, x_2, \cdots\} \cup \{x\}$ is compact.

Theorem 3.58. *Let (M, d) be a metric space, and A be a subset of M. Then A is compact if and only if A is sequentially compact.*

Proof. "$\Rightarrow$" Since a compact set must be totally bounded $\big($a finite sub-cover of $\big\{B(x,r)\,\big|\,x\in K\big\}$ suffices the purpose$\big)$, it suffices to show the completeness of K. Let $\{x_k\}_{k=1}^\infty$ be a Cauchy sequence in K. Suppose that $\{x_k\}_{k=1}^\infty$ does not converge in K. Then Proposition 2.56 and 2.58 imply that every point of K is not a cluster point of $\{x_k\}_{k=1}^\infty$; thus

$$\forall\, y\in K, \exists\, \delta_y>0 \ni \#\big\{k\in\mathbb{N}\,|\,x_k\in B(y,\delta_y)\big\}<\infty. \tag{3.3.2}$$

The collection $\big\{B(y,\delta_y)\big\}_{y\in K}$ then is an open cover of K; thus possesses a finite sub-cover $\big\{B\big(y_i,\delta_{y_i}\big)\big\}_{i=1}^N$. In particular, $\{x_k\}_{k=1}^\infty \subseteq \bigcup\limits_{i=1}^N B\big(y_i,\delta_{x_i}\big)$ or

$$\#\Big\{k\in\mathbb{N}\,\Big|\,x_k\in\bigcup_{i=1}^N B(y_i,\delta_{y_i})\Big\}=\infty$$

which contradicts to (3.3.2).

"$\Leftarrow$" Let $\big\{U_\alpha\big\}_{\alpha\in I}$ be an open cover of K.

Claim: there exists $r>0$ such that for each $x\in K$, $B(x,r)\subseteq U_\alpha$ for some $\alpha\in I$.

Proof of claim: Suppose the contrary that for each $k\in\mathbb{N}$, there exists $x_k\in K$ such that $B\big(x_k,\dfrac{1}{k}\big)\nsubseteq U_\alpha$ for all $\alpha\in I$. Then $\{x_k\}_{k=1}^\infty$ is a sequence in K; thus by the assumption of sequential compactness, there exists a convergent subsequence $\{x_{k_j}\}_{j=1}^\infty$ with limit $x\in K$. Since $\{U_\alpha\}_{\alpha\in I}$ is an open cover of K, $x\in U_\beta$ for some $\beta\in I$. Then

(1) there is $r>0$ such that $B(x,r)\subseteq U_\beta$ since U_β is open.

(2) there exists $N>0$ such that $d(x_{k_j},x)<\dfrac{r}{2}$ for all $j\geqslant N$.

Choose $j\geqslant N$ such that $\dfrac{1}{k_j}<\dfrac{r}{2}$. Then $B\big(x_{k_j},\dfrac{1}{k_j}\big)\subseteq B(x,r)\subseteq U_\beta$, a contradiction.

Having established the claim, by the fact that K is totally bounded (Theorem 3.53) there exists $\{y_1,\cdots,y_N\}\subseteq K$ such that $K\subseteq\bigcup\limits_{i=1}^N B(y_i,r)$. For each $1\leqslant i\leqslant N$, the claim above implies that there exists $\alpha_i\in I$ such that $B(y_i,r)\subseteq U_{\alpha_i}$. Then $\bigcup\limits_{i=1}^N B(y_i,r)\subseteq\bigcup\limits_{i=1}^N U_{\alpha_i}$ which implies that

$$K\subseteq\bigcup_{i=1}^N U_{\alpha_i}\,. \qquad \square$$

Remark 3.59. For a given open cover $\{U_\alpha\}_{\alpha\in I}$ of a compact set K, the positive number r appearing in the claim above is called a ***Lebesgue number*** for the open cover. It has the property that for each $x \in K$ there exists $\alpha \in I$ such that $B(x,r) \subseteq U_\alpha$.

Definition 3.60. Let (M,d) be a metric space. A subset A of M is called ***pre-compact*** if $\bar{A}$ is compact. Let $U \subseteq M$ be an open set, a subset A of U is said to be ***compactly embedded*** in U, denoted by $A \subset\subset U$, if A is pre-compact and $\bar{A} \subseteq U$.

Remark 3.61. Suppose that A is a pre-compact set in (M,d). If $\{x_k\}_{k=1}^\infty$ be a sequence in A, then $\{x_k\}_{k=1}^\infty$ is a sequence in $\bar{A}$; thus the (sequential) compactness of $\bar{A}$ implies that there exists convergent subsequence $\{x_{k_j}\}_{j=1}^\infty$ (with limit in $\bar{A}$). Therefore, every sequence in a pre-compact set has a convergent subsequence.

Example 3.62. Let (M,d) be a complete metric space, and $A \subseteq M$ be totally bounded. Then $\bar{A}$ is totally bounded (by enlarging the radius of the balls); thus Theorem 3.27 and 3.53 imply that $\bar{A}$ is sequentially compact. In other words, in a complete metric space, totally bounded sets are pre-compact.

Remark 3.63. A generalized version of the Bolzano-Weierstrass property in a general metric space is the following: a metric space is said to satisfy the Bolzano-Weierstrass property if every totally bounded sequence has a convergent subsequence. Then a metric space is complete if and only if it satisfies the Bolzano-Weierstrass property.

3.4 Connectedness（連通性）

Definition 3.64. Let (M,d) be a metric space, and A be a subset of M. Two non-empty open sets U and V are said to separate A if

1. $A \cap U \cap V = \varnothing$; 2. $A \cap U \neq \varnothing$; 3. $A \cap V \neq \varnothing$; 4. $A \subseteq U \cup V$.

We say that A is ***disconnected*** or ***separated*** if such separation exists, and A is ***connected*** if no such separation exists.

Proposition 3.65. *Let (M,d) be a metric space. A subset $A \subseteq M$ is disconnected if and only if $A = A_1 \cup A_2$ with $A_1 \cap \bar{A}_2 = \bar{A}_1 \cap A_2 = \varnothing$ for some non-empty A_1 and A_2.*

Proof. "$\Rightarrow$" Suppose that there exist U, V non-empty open sets such that 1-4 in Definition 3.64 hold. Let $A_1 = A \cap U$ and $A_2 = A \cap V$. By 1, $A_1 \subseteq V^\complement$; thus Theorem 3.5 implies that $\bar{A}_1 \subseteq V^\complement$. This shows that $\bar{A}_1 \cap A_2 = \varnothing$. Similarly, $\bar{A}_2 \cap A_1 = \varnothing$.

"$\Leftarrow$" Let $U = \text{cl}(A_2)^\complement$ and $V = \text{cl}(A_1)^\complement$ be two open sets. Then $V \cap A_1 = U \cap A_2 = \varnothing$; thus

$$A \cap U \cap V = (A_1 \cup A_2) \cap U \cap V = (A_1 \cap U) \cap V = U \cap (A_1 \cap V) = \varnothing .$$

Moreover, since $A_1 \cap \bar{A}_2 = A_2 \cap \bar{A}_1 = \varnothing$, $A_1 \subseteq \text{cl}(A_2)^\complement = U$ and $A_2 \subseteq \text{cl}(A_1)^\complement = V$ so that Property 2 and 3 in Definition 3.64 hold. Finally, since $\bar{A}_1 \subseteq A_2^\complement$ and $\bar{A}_2 \subseteq A_1^\complement$, we have

$$(U \cup V)^\complement = U^\complement \cap V^\complement = \bar{A}_2 \cap \bar{A}_1 \subseteq A_1^\complement \cap A_2^\complement = (A_1 \cup A_2)^\complement = A^\complement$$

which implies that $A \subseteq U \cup V$. Therefore, A is disconnected. □

Proposition 3.65 implies the following alternative definition of connected sets (without defining disconnected sets first):

Definition 3.66. Let (M, d) be a metric space. A subset A of M is said to be ***connected*** if A cannot be represented as the union of two non-empty disjoint sets neither of which contains a limit point of the other.

Corollary 3.67. *Let (M, d) be a metric space. Suppose that a subset $A \subseteq M$ is connected, and $A = A_1 \cup A_2$, where $A_1 \cap \bar{A}_2 = \bar{A}_1 \cap A_2 = \varnothing$. Then A_1 or A_2 is empty.*

Theorem 3.68. *A subset A of the Euclidean space $(\mathbb{R}, |\cdot|)$ is connected if and only if it has the property that if $x, y \in A$ and $x < z < y$, then $z \in A$.*

Proof. "$\Rightarrow$" Suppose that there exist $x, y \in A$, $x < z < y$ but $z \notin A$. Then $A = A_1 \cup A_2$, where

$$A_1 = A \cap (-\infty, z) \quad \text{and} \quad A_2 = A \cap (z, \infty) .$$

Since $x \in A_1$ and $y \in A_2$, A_1 and A_2 are non-empty. Moreover, $\bar{A}_1 \cap A_2 = A_1 \cap \bar{A}_2 = \varnothing$; thus by Proposition 3.65, A is disconnected, a contradiction.

"$\Leftarrow$" Suppose the contrary that A is not connected (disconnected). Then there exist non-empty sets A_1 and A_2 such that $A = A_1 \cup A_2$ with $\bar{A}_1 \cap A_2 = A_1 \cap \bar{A}_2 = \varnothing$. Pick $x \in A_1$ and $y \in A_2$. W.L.O.G., we may assume that $x < y$. Define $z = \sup(A_1 \cap [x, y])$.

Claim: $z \in \bar{A}_1$.

Proof of claim: By definition, for any $n > 0$ there exists $x_n \in A_1 \cap [x, y]$ such that $z - \dfrac{1}{n} < x_n \leqslant z$. Therefore, $x_n \to z$ as $n \to \infty$ which implies that $z \in \bar{A}_1$.
Since $z \in \bar{A}_1$, $z \notin A_2$. In particular, $x \leqslant z < y$.

(a) If $z \notin A_1$, then $x < z < y$ and $z \notin A$, a contradiction.

(b) If $z \in A_1$, then $z \notin \bar{A}_2$; thus there exists $0 < r < y - z$ such that $(z - r, z + r) \subseteq \mathrm{cl}(A_2)^{\complement}$. Then for all $z_1 \in (z, z + r)$, $z < z_1 < y$ and $z_1 \notin A_2$. Then $x < z_1 < y$ and $z_1 \notin A$, a contradiction. □

Corollary 3.69. *Connected sets in* $(\mathbb{R}, |\cdot|)$ *are intervals.*

3.5 Subspace Topology

Let (M, d) be a metric space, and $N \subseteq M$ be a subset. Then (N, d) is a metric space, and the topology of (N, d) is called the ***subspace topology*** of (N, d).

Remark 3.70. The topology of a metric space is the collection of all open sets of that metric space.

Proposition 3.71. *Let* (M, d) *be a metric space, and* N *be a subset of* M*. A subset* $V \subseteq N$ *is open in* (N, d) *if and only if* $V = U \cap N$ *for some open set* U *in* (M, d)*.*

Proof. Let $B_N(x, r)$ denote the r-ball about x in (N, d); that is,

$$B_N(x, r) \equiv \left\{ y \in N \,\middle|\, d(x, y) < r \right\} \subseteq V .$$

Note that $B_N(x, r) = B(x, r) \cap N$, where $B(x, r)$ is the r-ball about x in the metric space (M, d).

"$\Rightarrow$" Let $V \subseteq N$ be open in (N, d). Then for all $x \in M$, there exists $r_x > 0$ such that $B_N(x, r_x) \subseteq V$. In particular,

$$V = \bigcup_{x \in V} B_N(x, r_x) = \bigcup_{x \in V} B(x, r) \cap N .$$

Define $U = \bigcup\limits_{x \in V} B(x, r_x)$. Then U is open in (M, d) (by Proposition 3.33), and

$$V = \bigcup_{x \in V} B(x, r_x) \cap N = U \cap N .$$

"$\Leftarrow$" Suppose that $V = U \cap N$ for some open set U in (M, d). Let $x \in V$. Then $x \in U$; thus there exists $r > 0$ such that $B(x, r) \subseteq U$. Therefore,

$$B_N(x, r) = B(x, r) \cap N \subseteq U \cap N = V$$

which implies that x is an interior point of V. This shows that V is open in (N, d). □

Corollary 3.72. *Let (M, d) be a metric space, and N be a subset of M. A subset $E \subseteq N$ is closed in (N, d) if and only if $E = F \cap N$ for some closed set F in (M, d).*

Proof. Note that if A, B are subsets of N, then $A = B$ if and only if $N \cap A^{\complement} = N \cap B^{\complement}$, where $A^{\complement}$ denote the set $M \backslash A$. Then for a subset E of N,

$$\begin{aligned}
E \text{ is closed in } (N, d) &\Leftrightarrow N \backslash E \text{ is open in } (N, d) \Leftrightarrow N \cap E^{\complement} \text{ is open in } (M, d) \\
&\Leftrightarrow N \cap E^{\complement} = N \cap U \text{ for some open set } U \text{ in } (M, d) \\
&\Leftrightarrow N \cap (N \cap E^{\complement})^{\complement} = N \cap (N \cap U)^{\complement} \text{ for some open set } U \text{ in } (M, d) \\
&\Leftrightarrow N \cap E = N \cap U^{\complement} \text{ for some open set } U \text{ in } (M, d) \\
&\Leftrightarrow N \cap E = N \cap F \text{ for some closed set } F \text{ in } (M, d) \\
&\Leftrightarrow E = F \cap N \text{ for some closed set } F \text{ in } (M, d).
\end{aligned}$$

□

Remark 3.73. Let (M, d) be a metric space, N be a subset of M, and $\{x_n\}_{n=1}^{\infty}$ be a sequence in N. We note that the convergence of $\{x_n\}_{n=1}^{\infty}$ in (N, d) implies the convergence of $\{x_n\}_{n=1}^{\infty}$ in (M, d), but not vice versa. For example, the sequence $\left\{\frac{1}{n}\right\}_{n=1}^{\infty}$ is convergent in $(\mathbb{R}, |\cdot|)$ but is not convergent in $((0, \infty), d)$, where d is the metric induced from the norm $|\cdot|$. Since the concept of convergence is different in a subspace, we expect that in a subspace the concept of closed will be different. In other words, the concept of closedness (and openness as well) of sets highly depends on the background metric space.

Definition 3.74. Let (M, d) be a metric space, and N be a subset of M. A subset A is said to be $\begin{array}{c}\text{open}\\ \text{closed}\\ \text{compact}\end{array}$ relative to N if $A \cap N$ is $\begin{array}{c}\text{open}\\ \text{closed}\\ \text{compact}\end{array}$ in the metric space (N, d).

Note that base on the definition above, Proposition 3.71 and Corollary 3.72 imply that if A is closed/open in (M, d), then A is relative closed/open in (N, d). However, we note that if $A \subseteq N$ is closed/open in (N, d), A is not necessary closed/open in (M, d).

Example 3.75. Let (M, d) be $(\mathbb{R}, |\cdot|)$, and $N = \mathbb{Q}$. Consider the set $F = [0, 1] \cap \mathbb{Q}$. By Corollary 3.72 F is closed in $(\mathbb{Q}, |\cdot|)$; however, F is not closed in $(\mathbb{R}, |\cdot|)$ since it is not complete (a Cauchy sequence in F might not converge).

Theorem 3.76. *Let (M, d) be a metric space, and N be a subset of M. A subset A of M is closed relative to N if and only if $A^{\complement}$ is open relative to N.*

Proof. Note that if A, B are subsets of N, then $A = B$ if and only if $N \cap A^{\complement} = N \cap B^{\complement}$, where $A^{\complement}$ denote the set $M \backslash A$. Then for a subset A of M,

$$\begin{aligned} A \text{ is closed relative to } N &\Leftrightarrow A \cap N \text{ is closed in } (N, d) \\ &\Leftrightarrow A \cap N = F \cap N \text{ for some closed set } F \text{ in } (M, d) \\ &\Leftrightarrow N \cap (A \cap N)^{\complement} = N \cap (F \cap N)^{\complement} \text{ for some closed set } F \text{ in } (M, d) \\ &\Leftrightarrow N \cap A^{\complement} = N \cap F^{\complement} \text{ for some closed set } F \text{ in } (M, d) \\ &\Leftrightarrow N \cap A^{\complement} = N \cap U \text{ for some open set } U \text{ in } (M, d) \end{aligned}$$

which, by Proposition 3.71, implies that $A^{\complement}$ is open relative to N. □

Theorem 3.77. *Let (M, d) be a metric space, and $K \subseteq N \subseteq M$. Then K is compact in (M, d) if and only if K is compact in (N, d).*

Proof. "⇒" Let $\{V_\alpha\}_{\alpha \in I}$ be an open cover of K in (N, d). By Proposition 3.71 there exists a collection of open sets $\{U_\alpha\}_{\alpha \in I}$ in (M, d) such that $V_\alpha = U_\alpha \cap N$ for all $\alpha \in I$. Since $\{V_\alpha\}_{\alpha \in I}$ is an open cover of K in (N, d), $\{U_\alpha\}_{\alpha \in I}$ is a cover of K in (M, d); thus by the compactness of K in (M, d), there exists $\alpha_1, \cdots, \alpha_N$ such that

$$K \subseteq \bigcup_{j=1}^{N} U_{\alpha_j}.$$

Since $K \subseteq N$ and $V_\alpha = U_\alpha \cap N$, we must have $K \subseteq \bigcup_{j=1}^{N} V_{\alpha_j}$ which shows that there is a finite subcover of K for the open cover $\{V_\alpha\}_{\alpha \in I}$. Therefore, K is compact in (N, d).

"⇐" Let $\{U_\alpha\}_{\alpha \in I}$ be an open cover of K in (M, d). Define $V_\alpha = U_\alpha \cap N$. Since $K \subseteq N$, Proposition 3.71 implies that $\{V_\alpha\}_{\alpha \in I}$ is an open cover of K in (N, d); thus the compactness of K in (N, d) implies that there exists $\alpha_1, \cdots, \alpha_N$ such that

$$K \subseteq \bigcup_{j=1}^{N} V_{\alpha_j}.$$

Since $V_\alpha \subseteq U_\alpha$ for all $\alpha \in I$, $K \subseteq \bigcup_{j=1}^{N} U_{\alpha_j}$ which shows that there is a finite subcover of K for the open cover $\{U_\alpha\}_{\alpha \in I}$. Therefore, K is compact in (M, d). □

Alternative proof - sketch. Let $\{x_k\}_{k=1}^{\infty} \subseteq K$ be a sequence. By sequential compactness of K in either (M, d) or (N, d), there exists $\{x_{k_j}\}_{j=1}^{\infty}$ and $x \in K$ such that $x_{k_j} \to x$ as $j \to \infty$.

As long as the metric d used in different space are identical, the concept of convergence of a sequence are the same; thus (sequential) compactness in (M, d) or (N, d) are the same. □

Example 3.78. Let (M, d) be $(\mathbb{R}, |\cdot|)$, and $N = \mathbb{Q}$. Then $F = [0,1] \cap \mathbb{Q}$ is not compact in $(\mathbb{Q}, |\cdot|)$ since F is not complete. We can also apply Theorem 3.77 to see this: if $F \subseteq \mathbb{Q}$ is compact in $(\mathbb{Q}, |\cdot|)$, then F is compact in $(\mathbb{R}, |\cdot|)$ which is clearly not the case since F is not even closed in $(\mathbb{R}, |\cdot|)$.

Remark 3.79. Let (M, d) be a metric space. By Proposition 3.65 a subset $A \subseteq M$ is disconnected if and only if there exist two subsets U_1, U_2 of A, open relative to A, such that $A = U_1 \cup U_2$ and $U_1 \cap U_2 = \varnothing$ (one choice of (U_1, U_2) is $U_1 = A\backslash\bar{A}_1$ and $U_2 = A\backslash\bar{A}_2$, where A_1 and A_2 are given by Proposition 3.65). Note that U_1 and U_2 are also closed relative to A.

Given the observation above, if A is a connected set and E is a subset of A such that E is closed and open relative to A, then $E = \varnothing$ or $E = A$.

3.6 Exercises

In the exercise section of this chapter, we first introduce the concepts of accumulation points, isolated points and derived set of a set as follows.

Definition 3.80. Let (M, d) be a normed vector space, and A be a subset of M.

1. A point $x \in M$ is called an ***accumulation point*** of A if there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in $A\backslash\{x\}$ such that $\{x_n\}_{n=1}^{\infty}$ converges to x.

2. A point $x \in A$ is called an ***isolated point*** (孤立點) (of A) if there exists no sequence in $A\backslash\{x\}$ that converges to x.

3. The ***derived set*** of A is the collection of all accumulation points of A, and is denoted by A'.

§ 3.1 Limit Points and Interior Points of Sets

Problem 3.1. Let (M, d) be a metric space, and A be a subset of M.

1. Show that the collection of all isolated points of A is $A\backslash A'$.

2. Show that $A' = \bar{A} \backslash (A \backslash A')$. In other words, the derived set consists of all limit points that are not isolated points. Also show that $\bar{A} \backslash A' = A \backslash A'$.

Problem 3.2. Let A and B be subsets of a metric space (M, d). Show that

1. $\mathrm{cl}(\mathrm{cl}(A)) = \mathrm{cl}(A)$.

2. $\mathrm{cl}(A \cup B) = \mathrm{cl}(A) \cup \mathrm{cl}(B)$.

3. $\mathrm{cl}(A \cap B) \subseteq \mathrm{cl}(A) \cap \mathrm{cl}(B)$. Find examples of that $\mathrm{cl}(A \cap B) \subsetneq \mathrm{cl}(A) \cap \mathrm{cl}(B)$.

Problem 3.3. Let A and B be subsets of a metric space (M, d). Show that

1. $\mathrm{int}(\mathrm{int}(A)) = \mathrm{int}(A)$.

2. $\mathrm{int}(A \cap B) = \mathrm{int}(A) \cap \mathrm{int}(B)$.

3. $\mathrm{int}(A \cup B) \supseteq \mathrm{int}(A) \cup \mathrm{int}(B)$. Find examples of that $\mathrm{int}(A \cup B) \supsetneq \mathrm{int}(A) \cup \mathrm{int}(B)$.

Problem 3.4. Let (M, d) be a metric space, and A be a subset of M. Show that

$$\partial A = \big(A \cap \mathrm{cl}(M \backslash A)\big) \cup \big(\mathrm{cl}(A) \backslash A\big).$$

Problem 3.5. Recall that in a metric space (M, d), a subset A is said to be dense in S if subsets satisfy $A \subseteq S \subseteq \mathrm{cl}(A)$. For example, $\mathbb{Q}$ is dense in $\mathbb{R}$.

1. Show that if A is dense in S and if S is dense in T, then A is dense in T.

2. Show that if A is dense in S and $B \subseteq S$ is open, then $B \subseteq \mathrm{cl}(A \cap B)$.

Problem 3.6. Let A and B be subsets of a metric space (M, d). Show that

1. $\partial(\partial A) \subseteq \partial(A)$. Find examples of that $\partial(\partial A) \subsetneq \partial A$. Also show that $\partial(\partial A) = \partial A$ if A is closed.

2. $\partial(A \cup B) \subseteq \partial A \cup \partial B \subseteq \partial(A \cup B) \cup A \cup B$. Find examples of that equalities do not hold.

3. If $\mathrm{cl}(A) \cap \mathrm{cl}(B) = \varnothing$, then $\partial(A \cup B) = \partial A \cup \partial B$.

4. $\partial(A \cap B) \subseteq \partial A \cup \partial B$. Find examples of the equalities do not hold.

5. $\partial(\partial(\partial A)) = \partial(\partial A)$.

§ 3.2 Closed Sets and Open Sets

Problem 3.7. Let (M, d) be a metric space, and A be a subset of M. Show that $A \supseteq A'$ if and only if A is closed.

Problem 3.8. Show that the derived set of a set (in a metric space) is closed.

Problem 3.9. Let $A \subseteq \mathbb{R}^n$. Define the sequence of sets $A^{(m)}$ as follows: $A^{(0)} = A$ and $A^{(m+1)} =$ the derived set of $A^{(m)}$ for $m \in \mathbb{N}$. Complete the following.

1. Prove that each $A^{(m)}$ for $m \in \mathbb{N}$ is a closed set; thus $A^{(1)} \supseteq A^{(2)} \supseteq \cdots$.

2. Show that if there exists some $m \in \mathbb{N}$ such that $A^{(m)}$ is a countable set, then A is countable.

3. For any given $m \in \mathbb{N}$, is there a set A such that $A^{(m)} \neq \varnothing$ but $A^{(m+1)} = \varnothing$?

4. Let A be uncountable. Then each $A^{(m)}$ is an uncountable set. Is it possible that $\bigcap\limits_{m=1}^{\infty} A^{(m)} = \varnothing$?

5. Let $A = \left\{\dfrac{1}{m} + \dfrac{1}{k} \,\middle|\, m - 1 > k(k-1), m, k \in \mathbb{N}\right\}$. Find $A^{(1)}$, $A^{(2)}$ and $A^{(3)}$.

Problem 3.10. Recall that a cluster point x of a sequence $\{x_n\}_{n=1}^{\infty}$ satisfies that

$$\forall\, \varepsilon > 0, \#\big\{n \in \mathbb{N} \,\big|\, x_n \in B(x, \varepsilon)\big\} = \infty\,.$$

Show that the collection of cluster points of a sequence (in a metric space) is closed.

Problem 3.11. Let (M, d) be a metric space, and $A \subseteq M$. Show that A is dense in M if and only if $A \cap U \neq \varnothing$ for all non-empty open set $U \subseteq M$.

Problem 3.12. Determine whether the following sets are open or not.

1. $\bigcup\limits_{n=1}^{\infty} \Big[-2 + \dfrac{1}{n}, 2 + \dfrac{1}{n}\Big]$.

2. $\bigcup\limits_{n=1}^{\infty} \Big[-2 - \dfrac{1}{n}, 2 - \dfrac{1}{n}\Big]$.

3. $\bigcup\limits_{n=1}^{\infty} \Big[-2 + \dfrac{1}{n}, 2 - \dfrac{1}{n}\Big]$.

4. $\displaystyle\bigcup_{n=1}^{\infty}\Big[-2-\frac{1}{n},2+\frac{1}{n}\Big]$.

Problem 3.13. Let $(\mathcal{V},\|\cdot\|)$ ba a normed vector space, and C be a non-empty convex set in $\mathcal{V}$.

1. Show that $\bar{C}$ is convex.

2. Show that if $\boldsymbol{x}\in\mathring{C}$ and $\boldsymbol{y}\in\bar{C}$, then $(1-\lambda)\boldsymbol{x}+\lambda\boldsymbol{y}\in\mathring{C}$ for all $\lambda\in(0,1)$. This result is sometimes called the ***line segment principle***.

3. Show that $\mathring{C}$ is convex (you may need the conclusion in 2 to prove this).

4. Show that $\mathrm{cl}(\mathring{C})=\mathrm{cl}(C)$.

5. Show that $\mathrm{int}(\bar{C})=\mathrm{int}(C)$.

Hint: 2. Prove by contradiction.
3 and 4. Use the line segment principle.
5. Show that $\boldsymbol{x}\in\mathrm{int}(\bar{C})$ can be written as $(1-\lambda)\boldsymbol{y}+\lambda\boldsymbol{z}$ for some $\boldsymbol{y}\in\mathring{C}$ and $\boldsymbol{z}\in B(\boldsymbol{x},\varepsilon)\subseteq\bar{C}$.

Problem 3.14. Let $(\mathcal{V},\langle\cdot,\cdot\rangle)$ be a finite dimensional inner product space with induced norm $\|\cdot\|$, and C be a non-empty convex set in $\mathcal{V}$. Show that for each $\boldsymbol{x}\in\mathcal{V}$, there is a unique point $\boldsymbol{x}_C\in\bar{C}$ such that

$$\|\boldsymbol{x}-\boldsymbol{x}_C\|=\mathrm{dist}(\boldsymbol{x},C)=\inf\big\{\|\boldsymbol{x}-\boldsymbol{y}\|\,\big|\,\boldsymbol{y}\in C\big\}.$$

How about if $(\mathcal{V},\|\cdot\|)$ is a norm vector space whose norm $\|\cdot\|$ is not induced from an inner product?

Problem 3.15. In this problem we discuss the concept of supporting hyperplanes in an inner product spaces. Recall that a hyperplane in an n-dimensional vector space is an $(n-1)$-dimensional subspace.

Definition 3.81. Let $(\mathcal{V},\|\cdot\|)$ be a normed vector space, and S be a subset of $\mathcal{V}$. A hyperplane H is called a supporting hyperplane of S if

1. S is a subset of one of the two closed half-spaces bounded by H;

2. H contains at least one boundary point of S.

If $\boldsymbol{x} \in \partial S$, a supporting hyperplane of S at $\boldsymbol{x}$ is a supporting hyperplane of S containing $\boldsymbol{x}$.

In particular, if the norm is induced by an inner product $\langle \cdot, \cdot \rangle$, H is a supporting hyperplane of S at $\boldsymbol{x} \in \partial S$ if there exists a unit vector $\boldsymbol{n} \in \mathcal{V}$ such that $H = \big\{\boldsymbol{y} \,\big|\, \langle \boldsymbol{n}, \boldsymbol{y} - \boldsymbol{x} \rangle = 0\big\}$ and $S \subseteq \big\{\boldsymbol{y} \,\big|\, \langle \boldsymbol{n}, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant 0\big\}$.

Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ ba an inner product space, and C be a non-empty convex set in $\mathcal{V}$. Show that for each $\boldsymbol{x} \in \partial C$, there is a supporting hyperplane of C at $\boldsymbol{x}$.

Problem 3.16. Let $(\mathcal{V}, \|\cdot\|)$ be a normed vector space. Show that for all $\boldsymbol{x} \in \mathcal{V}$ and $r > 0$,

$$\operatorname{int}\big(B[\boldsymbol{x}, r]\big) = B(\boldsymbol{x}, r)\,.$$

Is the identity above true in general metric space?

Problem 3.17. Let $\mathcal{M}_{n\times n}$ denote the collection of all $n \times n$ square real matrices, and $(\mathcal{M}_{n\times n}, \|\cdot\|_{p,q})$ be a normed space with norm $\|\cdot\|_{p,q}$ given in Problem 2.8. Show that the set

$$\mathrm{GL}(n) \equiv \big\{A \in \mathcal{M}_{n\times n} \,\big|\, \det(A) \neq 0\big\}$$

is an open set in $\mathcal{M}_{n\times n}$. The set $\mathrm{GL}(n)$ is called the general linear group.

Problem 3.18. Show that every open set in $\mathbb{R}$ is the union of at most countable collection of disjoint open intervals; that is, if $U \subseteq \mathbb{R}$ is open, then

$$U = \bigcup_{k\in\mathcal{I}} (a_k, b_k)\,,$$

where $\mathcal{I}$ is countable, and $(a_k, b_k) \cap (a_\ell, b_\ell) = \varnothing$ if $k \neq \ell$.

Hint: For each point $x \in U$, define $L_x = \big\{y \in \mathbb{R} \,\big|\, (y, x) \subseteq U\big\}$ and $R_x = \big\{y \in \mathbb{R} \,\big|\, (x, y) \subseteq U\big\}$. Define $I_x = (\inf L_x, \sup R_x)$. Show that $I_x = I_y$ if $(x, y) \in U$ and if $(x, y) \nsubseteq U$ then $I_x \cap I_y = \varnothing$

Problem 3.19. Let (M, d) be a metric space. A set $A \subseteq M$ is said to be ***perfect*** if $A = A'$ (so that there is no isolated points). The Cantor set is constructed by the following procedure: let $E_0 = [0, 1]$. Remove the segment $\big(\frac{1}{3}, \frac{2}{3}\big)$, and let E_1 be the union of the intervals

$$\big[0, \frac{1}{3}\big], \big[\frac{2}{3}, 1\big].$$

Remove the middle thirds of these intervals, and let E_2 be the union of the intervals

$$\big[0, \frac{1}{9}\big], \big[\frac{2}{9}, \frac{3}{9}\big], \big[\frac{6}{9}, \frac{7}{9}\big], \big[\frac{8}{9}, 1\big].$$

Continuing in this way, we obtain a sequence of closed set E_k such that

(a) $E_1 \supseteq E_2 \supseteq E_2 \supseteq \cdots$;

(b) E_n is the union of 2^n intervals, each of length 3^{-n}.

The set $C = \bigcap_{n=1}^{\infty} E_n$ is called the ***Cantor set***.

1. Show that C is a perfect set.

2. Show that C is uncountable.

3. Find $\operatorname{int}(C)$.

Problem 3.20. In this problem you are asked to prove the Baire Category Theorem.

Theorem (Baire Category Theorem). *Let (X, d) be a complete metric space. If $U_1, U_2, \dots$ are dense open subsets of X, then*

$$\bigcap_{n=1}^{\infty} U_n$$

is dense in X.

Complete the following.

1. By Problem 3.11, it suffices to show that

$$W \cap \bigcap_{n=1}^{\infty} U_n \neq \varnothing$$

for every nonempty open set $W \subseteq X$. Let W be a nonempty open set. Use the fact that each U_n is dense and open to recursively construct closed balls

$$B[x_1, r_1] \supseteq B[x_2, r_2] \supseteq \cdots$$

such that

$$B[x_n, r_n] \subseteq U_n \cap B(x_{n-1}, r_{n-1}) \qquad \text{and} \qquad r_n < 2^{-n}.$$

2. Show that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence, and hence converges to some limit $x \in X$.

3. Show that

$$x \in \bigcap_{n=1}^{\infty} B[x_n, r_n],$$

and that x is the only point in this intersection.

4. Deduce that
$$x \in W \cap \bigcap_{n=1}^{\infty} U_n,$$
and therefore
$$W \cap \bigcap_{n=1}^{\infty} U_n \neq \varnothing.$$
Conclude that $\bigcap_{n=1}^{\infty} U_n$ is dense in X.

§ 3.3 Compactness

Problem 3.21. Let $\mathcal{V}$ be a vector fields over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\} \subseteq \mathcal{V}$ is a basis for $\mathcal{V}$; that is, every $\boldsymbol{x} \in \mathcal{V}$ can be uniquely expressed as
$$\boldsymbol{x} = x^{(1)}\mathbf{e}_1 + x^{(2)}\mathbf{e}_2 + \cdots + x^{(n)}\mathbf{e}_n = \sum_{i=1}^{n} x^{(i)}\mathbf{e}_i \,.$$
Define $\|\boldsymbol{x}\|_2 = \Big(\sum\limits_{i=1}^{n} \big|x^{(i)}\big|^2\Big)^{\frac{1}{2}}$.

1. Show that $\|\cdot\|_2$ is a norm on $\mathcal{V}$.
2. Show that K is compact in $(\mathcal{V}, \|\cdot\|_2)$ if and only if K is closed and bounded.

Problem 3.22. Let (M, d) be a metric space.

1. Show that a closed subset of a compact set is compact.
2. Show that the union of a finite number of sequentially compact subsets of M is compact.
3. Show that the intersection of an arbitrary collection of sequentially compact subsets of M is sequentially compact.

Problem 3.23. A metric space (M, d) is said to be ***separable*** if there is a countable subset A which is dense in M. Show that every sequentially compact set is separable.
Hint: Consider the total boundedness using balls with radius $\dfrac{1}{n}$ for $n \in \mathbb{N}$.

Problem 3.24. Given $\{a_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ a bounded sequence, define
$$A = \big\{x \in \mathbb{R} \,\big|\, \text{there exists a subsequence } \{a_{k_j}\}_{j=1}^{\infty} \text{ such that } \lim_{j\to\infty} a_{k_j} = x\big\}\,.$$
Show that A is a non-empty sequentially compact set in $\mathbb{R}$. Furthermore , $\limsup\limits_{k\to\infty} a_k = \sup A$ and $\liminf\limits_{k\to\infty} a_k = \inf A$.

Problem 3.25. Let (M, d) be a metric space.

1. Show that if M is complete and A is a totally bounded subset of M, then $\mathrm{cl}(A)$ is sequentially compact.

2. Show that M is complete if and only if every totally bounded sequence has a convergent subsequence.

Problem 3.26. Let $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$d(x, y) = \begin{cases} |x_1 - y_1| & \text{if } x_2 = y_2\,, \\ |x_1 - y_1| + |x_2 - y_2| + 1 & \text{if } x_2 \neq y_2\,. \end{cases} \quad \text{where } x = (x_1, x_2) \text{ and } y = (y_1, y_2).$$

Problem 2.15 shows that d is a metric on $\mathbb{R}^2$. Consider the metric space $(\mathbb{R}^2, d)$.

1. Find $B(x, r)$ with $r < 1$, $r = 1$ and $r > 1$.

2. Show that the set $\{c\} \times [a, b] \subseteq (\mathbb{R}^2, d)$ is closed and bounded.

3. Examine whether the set $\{c\} \times [a, b] \subseteq (\mathbb{R}^2, d)$ is sequentially compact or not.

Problem 3.27. Let $\{x_k\}_{k=1}^{\infty}$ be a convergent sequence in a metric space, and $x_k \to x$ as $k \to \infty$. Show that the set $A \equiv \{x_1, x_2, \cdots, \} \cup \{x\}$ is sequentially compact.

Problem 3.28. 1. Show the so-called "***Finite Intersection Property***":

> Let (M, d) be a metric space, and K be a subset of M. Then K is compact if and only if for any family of closed subsets $\{F_\alpha\}_{\alpha \in I}$, we have
>
> $$K \cap \bigcap_{\alpha \in I} F_\alpha \neq \varnothing$$
>
> whenever $K \cap \bigcap\limits_{\alpha \in J} F_\alpha \neq \varnothing$ for all $J \subseteq I$ satisfying $\#J < \infty$.

2. Show the so-called "***Nested Set Properpty***":

> Let (M, d) be a metric space. If $\{K_n\}_{n=1}^{\infty}$ is a sequence of non-empty compact sets in M such that $K_j \supseteq K_{j+1}$ for all $j \in \mathbb{N}$, then there exists at least one point in $\bigcap\limits_{j=1}^{\infty} K_j$; that is,
>
> $$\bigcap_{j=1}^{\infty} K_j \neq \varnothing\,.$$

Problem 3.29. Let (M, d) be a metric space, and M itself is a sequentially compact set. Show that if $\{F_k\}_{k=1}^{\infty}$ is a sequence of closed sets such that $\mathrm{int}(F_k) = \varnothing$, then $M \setminus \bigcup_{k=1}^{\infty} F_k \neq \varnothing$.

Problem 3.30. Let ℓ^2 be the collection of all sequences $\{x_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ such that $\sum_{k=1}^{\infty} |x_k|^2 < \infty$. In other words,

$$\ell^2 = \Big\{\{x_k\}_{k=1}^{\infty} \,\Big|\, x_k \in \mathbb{R} \text{ for all } k \in \mathbb{N}, \sum_{k=1}^{\infty} |x_k|^2 < \infty\Big\}.$$

Define $\|\cdot\|_2 : \ell^2 \to \mathbb{R}$ by

$$\big\|\{x_k\}_{k=1}^{\infty}\big\|_2 = \Big(\sum_{k=1}^{\infty} |x_k|^2\Big)^{\frac{1}{2}}.$$

1. Show that $\|\cdot\|_2$ is a norm on ℓ^2. The normed space $(\ell^2, \|\cdot\|)$ usually is denoted by ℓ^2.
2. Show that $\|\cdot\|_2$ is induced by an inner product.
3. Show that $(\ell^2, \|\cdot\|_2)$ is complete.
4. Let $A = \big\{\boldsymbol{x} \in \ell^2 \,\big|\, \|\boldsymbol{x}\|_2 \leqslant 1\big\}$. Is A sequentially compact or not?

Problem 3.31. Let A, B be two non-empty subsets in $\mathbb{R}^n$. Define

$$d(A, B) = \inf\big\{\|x - y\|_2 \,\big|\, x \in A, y \in B\big\}$$

to be the distance between A and B. When $A = \{x\}$ is a point, we write $d(A, B)$ as $d(x, B)$ (which is consistent with the one given in Proposition 3.6).

(1) Prove that $d(A, B) = \inf\big\{d(x, B) \,\big|\, x \in A\big\}$.

(2) Show that $\big|d(x_1, B) - d(x_2, B)\big| \leqslant \|x_1 - x_2\|_2$ for all $x_1, x_2 \in \mathbb{R}^n$.

(3) Define $B_\varepsilon = \big\{x \in \mathbb{R}^n \,\big|\, d(x, B) < \varepsilon\big\}$ be the collection of all points whose distance from B is less than ε. Show that B_ε is open and $\bigcap_{\varepsilon>0} B_\varepsilon = \mathrm{cl}(B)$.

(4) If A is sequentially compact, show that there exists $x \in A$ such that $d(A, B) = d(x, B)$.

(5) If A is closed and B is sequentially compact, show that there exists $x \in A$ and $y \in B$ such that $d(A, B) = d(x, y)$.

(6) If A and B are both closed, does the conclusion of (5) hold?

Problem 3.32. Let $\mathcal{K}(n)$ denote the collection of all non-empty sequentially compact sets in $\mathbb{R}^n$. Define the Hausdorff distance of $K_1, K_2 \in \mathcal{K}(n)$ by

$$d^H(K_1, K_2) = \max\Big\{ \sup_{x \in K_2} d(x, K_1), \sup_{x \in K_1} d(x, K_2) \Big\},$$

in which $d(x, K)$ is the distance between x and K given in Problem 3.31. Show that $(\mathcal{K}(n), d^H)$ is a metric space.

Problem 3.33. Let $M = \big\{(x, y) \in \mathbb{R}^2 \,\big|\, x^2 + y^2 \leqslant 1\big\}$ with the standard metric $\|\cdot\|_2$. Show that $A \subseteq M$ is sequentially compact if and only if A is closed.

Problem 3.34. 1. Let $\{x_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ be a sequence in $(\mathbb{R}, |\cdot|)$ that converges to x and let $A_k = \{x_k, x_{k+1}, \cdots\}$. Show that $\{x\} = \bigcap\limits_{k=1}^{\infty} \overline{A_k}$. Is this true in any metric space?

2. Suppose that $\{K_j\}_{j=1}^{\infty}$ is a sequence of comapct non-empty sets satisfying the nested set property; that is, $K_j \supseteq K_{j+1}$, and $\operatorname{diam}(K_j) \to 0$ as $j \to \infty$, where

$$\operatorname{diam}(K_j) = \sup\big\{d(x, y) \,\big|\, x, y \in K_j\big\}.$$

Show that there is exactly one point in $\bigcap\limits_{j=1}^{\infty} K_j$.

Problem 3.35. Let (M, d) be a metric space, and A be a subset of M satisfying that every sequence in A has a convergent subsequence (with limit in M). Show that A is pre-compact. Remark: Together with Remark 3.61, we conclude that a subset A is pre-compact if and only if A has the property that "every sequence in A has a convergent subsequence".

§ 3.4 Connectedness

Problem 3.36. Let (M, d) be a metric space, and $A \subseteq M$. Show that A is disconnected (not connected) if and only if there exist non-empty closed set F_1 and F_2 such that

1. $A \cap F_1 \cap F_2 = \varnothing$; 2. $A \cap F_1 \neq \varnothing$; 3. $A \cap F_2 \neq \varnothing$; 4. $A \subseteq F_1 \cup F_2$.

Problem 3.37. Prove that if A is connected in a metric space (M, d) and $A \subseteq B \subseteq \bar{A}$, then B is connected.

Problem 3.38. Let (M,d) be a metric space, and $A \subseteq M$ be a subset. Suppose that A is connected and contain more than one point. Show that $A \subseteq A'$.

Problem 3.39. Show that the Cantor set C defined in Problem 3.19 is totally disconnected; that is, if $x, y \in C$, and $x \neq y$, then $x \in U$ and $y \in V$ for some open sets U, V separate C.

Problem 3.40. Let F_k be a nest of connected compact sets (that is, $F_{k+1} \subseteq F_k$ and F_k is connected for all $k \in \mathbb{N}$). Show that $\bigcap\limits_{k=1}^{\infty} F_k$ is connected. Give an example to show that compactness is an essential condition and we cannot just assume that F_k is a nest of closed connected sets.

Problem 3.41. Let $\{A_k\}_{k=1}^{\infty}$ be a family of connected subsets of M, and suppose that A is a connected subset of M such that $A_k \cap A \neq \varnothing$ for all $k \in \mathbb{N}$. Show that the union $\Big(\bigcup\limits_{k\in\mathbb{N}} A_k\Big) \cup A$ is also connected.

Problem 3.42. Let $A, B \subseteq M$ and A is connected. Suppose that $A \cap B \neq \varnothing$ and $A \cap B^{\complement} \neq \varnothing$. Show that $A \cap \partial B \neq \varnothing$.

Problem 3.43. Let (M,d) be a metric space and A be a non-empty subset of M. A maximal connected subset of A is called a ***connected component*** of A.

1. Let $a \in A$. Show that there is a unique connected components of A containing a.

2. Show that any two distinct connected components of A are disjoint. Therefore, A is the disjoint union of its connected components.

3. Show that every connected component of A is a closed subset of A.

4. If A is open, prove that every connected component of A is also open. Therefore, when $M = \mathbb{R}^n$, show that A has at most countable infinite connected components.

5. Find the connected components of the set of rational numbers or the set of irrational numbers in $\mathbb{R}$.

§ 3.6 Chapter Review

Problem 3.44 (True or False). Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. Every open set in a metric space is a countable union of closed sets.

2. Let $A \subseteq \mathbb{R}$ be bounded from above, then $\sup A \in A'$.

3. An infinite union of distinct closed sets cannot be closed.

4. An interior point of a subset A of a metric space (M, d) is an accumulation point of that set.

5. Let (M, d) be a metric space, and $A \subseteq M$. Then $(A')' = A'$.

6. There exists a metric space in which some unbounded Cauchy sequence exists.

7. Every metric defined in $\mathbb{R}^n$ is induced from some "norm" in $\mathbb{R}^n$.

8. There exists a non-zero dimensional normed vector space in which some compact non-zero dimensional linear subspace exists.

9. There exists a set $A \subseteq (0, 1]$ which is compact in $(0, 1]$ (in the sense of subspace topology), but A is not compact in $\mathbb{R}$.

10. Let $A \subseteq \mathbb{R}^n$ be a non-empty set. Then a subset B of A is compact in A if and only if B is closed and bounded in A.

Problem 3.45. Let (M, d) be a metric space, and $A \subseteq M$ be a subset. Determine which of the following statements are true.

1. $\text{int}A = A \backslash \partial A$.

2. $\text{cl}(A) = M \backslash \text{int}(M \backslash A)$.

3. $\text{int}(\text{cl}(A)) = \text{int}(A)$.

4. $\text{cl}(\text{int}(A)) = A$.

5. $\partial(\text{cl}(A)) = \partial A$.

6. If A is open, then $\partial A \subseteq M \backslash A$.

7. If A is open, then $A = \text{cl}(A) \backslash \partial A$. How about if A is not open?

Chapter 4

Continuous Maps

4.1 Continuity

Definition 4.1. Let (M, d) be a metric space, and A be a subset of M.

1. A point $x \in M$ is called an ***accumulation point*** of A if there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in $A \backslash \{x\}$ such that $\{x_n\}_{n=1}^{\infty}$ converges to x.
2. The ***derived set*** of A is the collection of all accumulation points of A, and is denoted by A'.

Remark 4.2. 1. A point $x \in A'$ if and only if $(\forall\, \varepsilon > 0)\big(B(x, \varepsilon) \cap (A \backslash \{x\}) \neq \varnothing\big)$. Therefore, $x \notin A'$ if and only if $(\exists\, \varepsilon > 0)\big(B(x, \varepsilon) \cap A \subseteq \{x\}\big)$.

2. A point $x \in M$ is an accumulation point x of A if and only if

$$(\forall\, \varepsilon > 0)\big(\#\big\{y \in M \,\big|\, y \in B(x, \varepsilon) \cap A\big\} = \infty\big).$$

Therefore, accumulation points of a set can be viewed as a generalization of cluster points of a sequence.

3. A subset A of M is closed if and only if $A \supseteq A'$. In fact, $\bar{A} = A \cup A'$.
4. The derived set A' of a subset A of M is closed.
5. A point $x \in A \backslash A'$ satisfies that there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \cap A = \{x\}$. Such kind of points are called ***isolated points*** of A.

Definition 4.3. Let (M, d) and (N, ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a map. For a given $c \in A'$, we say that the limit of f at c exists if for every sequence $\{x_k\}_{k=1}^{\infty} \subseteq A \backslash \{c\}$ converging to c, the sequence $\big\{f(x_k)\big\}_{k=1}^{\infty}$ converges（所有在定義域中取值不是 c 但收斂到 c 的數列，其函數值所形成的數列收斂）.

Similar to Proposition 2.42, the limit of f at x_0, if it exists, is unique.

Proposition 4.4. *Let (M,d) and (N,ρ) be metric spaces, A be a subset of M, $c \in A'$, and $f : A \to N$ be a map. If the limit of f at c exists, then the limit is unique in the sense that if $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ are sequences in $A\backslash\{c\}$ and $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} y_n = c$, then $\lim\limits_{n\to\infty} f(x_n) = \lim\limits_{n\to\infty} f(y_n)$.*

Proof. Let $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ be sequences in $A\backslash\{c\}$ so that $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} y_n = c$. Then $\lim\limits_{n\to\infty} f(x_n) = a$ and $\lim\limits_{n\to\infty} f(y_n) = b$ both exist. Define a new sequence

$$z_n = \begin{cases} x_{\frac{n+1}{2}} & \text{if } n \text{ is odd}, \\ y_{\frac{n}{2}} & \text{if } n \text{ is even}, \end{cases}$$

or $\{z_n\}_{n=1}^{\infty} = \{x_1, y_1, x_2, y_2, \cdots\}$. Then $z_n \to c$ as $n \to \infty$; thus $\lim\limits_{n\to\infty} f(z_n)$ exists. Since $\big\{f(x_n)\big\}_{n=1}^{\infty}$ and $\big\{f(y_n)\big\}_{n=1}^{\infty}$ are both subsequences of $\big\{f(z_n)\big\}_{n=1}^{\infty}$, by Proposition 2.56 we conclude that $a = b$. □

Notation: When the limit of f at c exists, we use $\lim\limits_{x\to c} f(x)$ to denote the common limit of $\lim\limits_{k\to\infty} f(x_k)$ if $\{x_k\}_{k=1}^{\infty} A\backslash\{c\}$ converges to c.

Proposition 4.5. *Let (M,d) and (N,ρ) be metric spaces, A be a subset of M, $c \in A'$, and $f : A \to N$ be a map. Then $\lim\limits_{x\to c} f(x) = b$ if and only if*

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)\big(0 < d(x,c) < \delta \text{ and } x \in A \Rightarrow \rho(f(x), b) < \varepsilon\big).$$

Proof. "$\Rightarrow$" Assume the contrary that there exists $\varepsilon > 0$ such that for all $\delta > 0$, there exists $x_\delta \in A$ with

$$0 < d(x_\delta, c) < \delta \quad \text{and} \quad \rho(f(x_\delta), b) \geqslant \varepsilon.$$

In particular, for each $k \in \mathbb{N}$, we can find $x_k \in A\backslash\{c\}$ such that

$$0 < d(x_k, c) < \frac{1}{k} \quad \text{and} \quad \rho(f(x_k), b) \geqslant \varepsilon.$$

Then $x_k \to c$ as $k \to \infty$ but $f(x_k) \not\to b$ as $k \to \infty$, a contradiction.

"$\Leftarrow$" Let $\{x_k\}_{k=1}^{\infty} \subseteq A\backslash\{c\}$ be such that $x_k \to c$ as $k \to \infty$, and $\varepsilon > 0$ be given. By assumption,

$$\exists\, \delta > 0 \ni \rho(f(x), b) < \varepsilon \quad \text{whenever} \quad 0 < d(x,c) < \delta \text{ and } x \in A.$$

Since $x_k \to c$ as $k \to \infty$, there exists $N > 0$ such that $d(x_k, c) < \delta$ if $k \geqslant N$. Therefore,

$$\rho(f(x_k), b) < \varepsilon \quad \forall\, k \geqslant N$$

which implies that $\lim\limits_{k\to\infty} f(x_k) = b$. □

Remark 4.6. The positive number δ in the proposition above usually depends on ε, as well as the point c. Therefore, we also write $\delta = \delta(c, \varepsilon)$ to emphasize the dependence of c and ε.

Remark 4.7. Let $(M, d) = (N, \rho) = (\mathbb{R}, |\cdot|)$, $A = (a, b)$, and $f : A \to N$. We write $\lim\limits_{x\to a^+} f(x)$ and $\lim\limits_{x\to b^-} f(x)$ for the limit $\lim\limits_{x\to a} f(x)$ and $\lim\limits_{x\to b} f(x)$, respectively, if the later exist.

Definition 4.8. Let (M, d) and (N, ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a map. For a given $c \in A$, f is said to be continuous at c if either $c \in A \backslash A'$ or $\lim\limits_{x\to c} f(x) = f(c)$.

Proposition 4.9. *Let (M, d) and (N, ρ) be metric spaces, A be a subset of M, $c \in A$ and $f : A \to N$ be a map. Then the following three statements are equivalent.*

1. *f is continuous at c.*
2. *For every convergent sequence $\{x_n\}_{n=1}^{\infty} \subseteq A$ with limit c, $\lim\limits_{n\to\infty} f(x_n) = f(c)$.*
3. *For each $\varepsilon > 0$, there exists $\delta = \delta(c, \varepsilon) > 0$ such that*

$$\rho(f(x), f(c)) < \varepsilon \quad \textit{whenever} \quad x \in B_M(c, \delta) \cap A.$$

In logical notation,

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)\big(x \in B_M(c, \delta) \cap A \Rightarrow f(x) \in B_N(f(c), \varepsilon)\big),$$

where $B_M(\cdot, \cdot)$ and $B_N(\cdot, \cdot)$ denote balls in (M, d) and (N, ρ), respectively.

Proof. "1 ⇒ 3" Note that $A = (A \cap A') \cup (A \backslash A')$.

Case 1: If $c \in A \cap A'$, then f is continuous at c if and only if

$$\forall\, \varepsilon > 0, \exists\, \delta = \delta(c, \varepsilon) > 0 \ni \rho(f(x), f(c)) < \varepsilon \;\text{ whenever }\; x \in B_M(c, \delta) \cap A \backslash \{c\}.$$

Since $\rho(f(c), f(c)) = 0 < \varepsilon$, we find that the statement above is equivalent to that

$$\forall\, \varepsilon > 0, \exists\, \delta = \delta(c, \varepsilon) > 0 \ni \rho(f(x), f(c)) < \varepsilon \quad \text{whenever} \quad x \in B_M(c, \delta) \cap A.$$

Case 2: Let $c \in A \backslash A'$. Then there exists $\delta > 0$ such that $B_M(c,\delta) \cap A = \{c\}$. Therefore,

$$x \in B_M(c,\delta) \cap A \Rightarrow \rho(f(x), f(c)) < \varepsilon$$

no matter what $\varepsilon > 0$ is given.

"$3 \Rightarrow 2$" The proof of this direction is almost identical as the proof of the direction "$\Leftarrow$" of Proposition 4.5. Let $\{x_k\}_{k=1}^{\infty} \subseteq A\backslash$ be such that $x_k \to c$ as $k \to \infty$, and $\varepsilon > 0$ be given. By assumption,

$$\exists\, \delta > 0 \ni \rho(f(x), b) < \varepsilon \quad \text{whenever} \quad d(x,c) < \delta \text{ and } x \in A\,.$$

Since $x_k \to c$ as $k \to \infty$, there exists $N > 0$ such that $d(x_k, c) < \delta$ if $k \geqslant N$. Therefore,

$$\rho(f(x_k), b) < \varepsilon \quad \forall\, k \geqslant N$$

which implies that $\lim\limits_{k \to \infty} f(x_k) = f(c)$.

"$2 \Rightarrow 1$" If $c \in A\backslash A'$, f is continuous at c; thus it suffices to show that f is continuous at c $\big($or equivalently, $\lim\limits_{x \to c} f(x) = f(c)\big)$ for $c \in A \cap A'$ when 2 is true.

Let $\{x_k\}_{k=1}^{\infty} \subseteq A\backslash\{c\}$ be a sequence with limit c. By assumption, $\lim\limits_{k\to\infty} f(x_k) = f(c)$; thus we establish that for every convergent sequence $\{x_k\}_{k=1}^{\infty} \subseteq A\backslash\{c\}$ with limit c, the sequence $\{f(x_k)\}_{k=1}^{\infty}$ converges to $f(c)$; thus $\lim\limits_{x\to c} f(x) = f(c)$. □

Remark 4.10. We remark here that Proposition 4.9 implies that f is continuous at $c \in A$ if and only if

$$\forall\, \varepsilon > 0, \exists\, \delta > 0 \ni f(B(c,\delta) \cap A) \subseteq B(f(c), \varepsilon)\,.$$

Example 4.11. Let $X = \mathscr{C}([a,b];\mathbb{R})$, the collection of all real-valued continuous functions defined on $[a,b]$, and $\|\cdot\|_X$ be the norm given by $\|f\|_X = \max\limits_{x\in[a,b]} \big|f(x)\big|$. Note that $(X, \|\cdot\|_X)$ is a normed vector space (Example 2.21). Define $\mathrm{I} : X \to \mathbb{R}$ by

$$\mathrm{I}(f) = \int_a^b \big|f(x)\big|^2\, dx\,.$$

In the following we show that I is continuous at any points on X. Let $f \in X$ and $\varepsilon > 0$ be given. Choose $\delta = \min\Big\{\varepsilon, \dfrac{\varepsilon}{2(b-a)\big(2\|f\|_X + \varepsilon\big)}\Big\}$. Then $0 < \delta \leqslant \varepsilon$ and if $g \in X$ satisfies

$\|f-g\|_X < \delta$, we must have

$$\begin{aligned}(b-a)\big[2\|f\|_X + \|f-g\|_X\big]\|f-g\|_X &\leqslant (b-a)\big(2\|f\|_X+\delta\big)\delta \\ &\leqslant (b-a)\big(2\|f\|_X+\varepsilon\big)\frac{\varepsilon}{2(b-a)\big(2\|f\|_X+\varepsilon\big)} = \frac{\varepsilon}{2} < \varepsilon .\end{aligned}$$

Therefore, if $g \in X$ and $\|g-f\|_X < \delta$,

$$\begin{aligned}\big|\mathrm{I}(g)-\mathrm{I}(f)\big| = \Big|\int_a^b \big[|g(x)|^2 - |f(x)|^2\big]\,dx\Big| &\leqslant \int_a^b \big|g(x)-f(x)\big|\big|g(x)+f(x)\big|\,dx \\ \leqslant (b-a)\big(\|f\|_X+\|g\|_X\big)\|f-g\|_X &\leqslant (b-a)\big(2\|f\|_X+\|f-g\|_X\big)\|f-g\|_X < \varepsilon;\end{aligned}$$

thus I is continuous on X.

Definition 4.12. Let (M,d) and (N,ρ) be metric spaces, and A be a subset of M. A map $f : A \to N$ is said to be continuous on the set $B \subseteq A$ if f is continuous at each point of B.

Remark 4.13. The Dirichlet function $f : [0,1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & \text{if } x \in [0,1] \cap \mathbb{Q}, \\ 1 & \text{if } x \in [0,1] \cap \mathbb{Q}^{\complement}. \end{cases}$$

is not continuous at any point of $[0,1]$; however, the restriction of f to $B = [0,1] \cap \mathbb{Q}$ (or $B = [0,1] \cap \mathbb{Q}^{\complement}$), denoted by $f\restriction_B$, is continuous on B. Therefore, f is continuous on B is different from that $f\restriction_B$ is continuous on B.

Theorem 4.14. *Let (M,d) and (N,ρ) be metric spaces, $A \subseteq M$, and $f : A \to N$ be a map. Then the following assertions are equivalent:*

1. *f is continuous on A.*
2. *For each open set $V \subseteq N$, $f^{-1}(V) \subseteq A$ is open relative to A; that is, $f^{-1}(V) = U \cap A$ for some U open in M.*
3. *For each closed set $E \subseteq N$, $f^{-1}(E) \subseteq A$ is closed relative to A; that is, $f^{-1}(E) = F \cap A$ for some F closed in M.*

Proof. It should be clear that $2 \Leftrightarrow 3$ (left as an exercise); thus we show that $1 \Leftrightarrow 2$. Before proceeding, we recall that $B \subseteq f^{-1}(f(B))$ for all $B \subseteq A$ and $f(f^{-1}(B)) \subseteq B$ for all $B \subseteq N$.

"1 ⇒ 2" Let $a \in f^{-1}(V)$. Then $f(a) \in V$. Since V is open in (N,ρ), there exists $\varepsilon_{f(a)} > 0$ such that $B_N(f(a), \varepsilon_{f(a)}) \subseteq V$. By continuity of f (and Remark 4.10), there exists $\delta_a > 0$ such that

$$f(B_M(a,\delta_a) \cap A) \subseteq B_N\big(f(a), \varepsilon_{f(a)}\big).$$

Therefore, by Proposition 0.11, for each $a \in f^{-1}(V)$, there exists $\delta_a > 0$ such that

$$B_M(a,\delta_a) \cap A \subseteq f^{-1}\big(f(B_M(a,\delta_a) \cap A)\big) \subseteq f^{-1}\big(B_N\big(f(a),\varepsilon_{f(a)}\big)\big) \subseteq f^{-1}(V). \quad (4.1.1)$$

Let $U = \bigcup\limits_{a \in f^{-1}(V)} B_M(a,\delta_a)$. Then U is open (since it is the union of arbitrarily many open balls), and

(a) $U \supseteq f^{-1}(V)$ since U contains every center of balls whose union forms U;

(b) $U \cap A \subseteq f^{-1}(V)$ by (4.1.1).

Therefore, $U \cap A = f^{-1}(V)$.

"2 ⇒ 1" Let $a \in A$ and $\varepsilon > 0$ be given. Define $V = B_N(f(a), \varepsilon)$. By assumption there exists an open set U in (M, d) such that $f^{-1}(V) = U \cap A$. Since $a \in f^{-1}(V)$, $a \in U$; thus by the openness of U, there exists $\delta > 0$ such that $B_M(a,\delta) \subseteq U$. Therefore, by Proposition 0.11 we have

$$f(B_M(a,\delta) \cap A) \subseteq f(U \cap A) = f(f^{-1}(V)) \subseteq V = B_N(f(a),\varepsilon)$$

which implies that f is continuous at a. Therefore, f is continuous at a for all $a \in A$; thus f is continuous on A. □

Example 4.15. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be continuous. Then $\big\{x \in \mathbb{R}^n \,\big|\, \|f(x)\|_2 < 1\big\}$ is open since

$$\big\{x \in \mathbb{R}^n \,\big|\, \|f(x)\|_2 < 1\big\} = f^{-1}\big(B(0,1)\big).$$

Example 4.16. Let $f : \mathcal{M}_{n\times n} \to \mathbb{R}$ be defined by $f(A) = \det(A)$. Then the set

$$\mathrm{GL}(n) \equiv \big\{A \in \mathcal{M}_{n\times n} \,\big|\, \det(A) \neq 0\big\}$$

is open in $(\mathcal{M}_{n\times n}, \|\cdot\|_{p,q})$ for all $1 \leqslant p, q \leqslant \infty$ (if one can show that the determinant function is continuous).

Remark 4.17. For a function f of two variables or more, it is important to distinguish the continuity of f and the continuity in each variable (by holding all other variables fixed). For example, let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} 1 & \text{if either } x = 0 \text{ or } y = 0, \\ 0 & \text{if } x \neq 0 \text{ and } y \neq 0. \end{cases}$$

Observe that $f(0,0) = 1$, but f is not continuous at $(0,0)$. In fact, for any $\delta > 0$, $f(x,y) = 0$ for infinitely many values of $(x,y) \in B((0,0),\delta)$; that is, $|f(x,y) - f(0,0)| = 1$ for such values. However if we consider the function $g(x) = f(x,0) = 1$ or the function $h(y) = f(0,y) = 1$, then g, h are continuous. Note also that $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ does not exists but $\lim\limits_{x\to 0}(\lim\limits_{y\to 0} f(x,y)) = \lim\limits_{x\to 0} 0 = 0$.

4.2 Operations on Continuous Maps

Definition 4.18. Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space over field $\mathbb{F}$, A be a subset of M, and $f, g : A \to \mathcal{V}$ be maps, $h : A \to \mathbb{F}$ be a function. The maps $f+g$, $f-g$ and hf, mapping from A to $\mathcal{V}$, are defined by

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \qquad && \forall\, x \in A, \\ (f-g)(x) &= f(x) - g(x) \qquad && \forall\, x \in A, \\ (hf)(x) &= h(x)f(x) \qquad && \forall\, x \in A. \end{aligned}$$

The map $\dfrac{f}{h} : A \backslash \{x \in A \,|\, h(x) = 0\} \to \mathcal{V}$ is defined by

$$\Big(\frac{f}{h}\Big)(x) = \frac{f(x)}{h(x)} \qquad \forall\, x \in A\backslash\{x \in A \,|\, h(x) = 0\}.$$

Proposition 4.19. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space over field $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$), A be a subset of M, and $f, g : A \to \mathcal{V}$ be maps, $h : A \to \mathbb{F}$ be a function. Suppose that $x_0 \in A'$, and $\lim\limits_{x\to x_0} f(x) = a$, $\lim\limits_{x\to x_0} g(x) = b$, $\lim\limits_{x\to x_0} h(x) = c$. Then*

$$\begin{aligned} \lim_{x\to x_0} (f+g)(x) &= a + b, \\ \lim_{x\to x_0} (f-g)(x) &= a - b, \\ \lim_{x\to x_0} (hf)(x) &= ca, \\ \lim_{x\to x_0} \Big(\frac{f}{h}\Big) &= \frac{a}{c} \quad \textit{if } c \neq 0. \end{aligned}$$

Corollary 4.20. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space over field $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$), A be a subset of M, and $f, g : A \to \mathcal{V}$ be maps, $h : A \to \mathbb{F}$ be a function. Suppose that f, g, h are continuous at $x_0 \in A$. Then the maps $f+g$, $f-g$ and hf are continuous at x_0, and $\dfrac{f}{h}$ is continuous at x_0 if $h(x_0) \neq 0$.*

Corollary 4.21. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space over field $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$), $A \subseteq M$, and $f, g : A \to \mathcal{V}$ be continuous maps, $h : A \to \mathbb{F}$ be a continuous function. Then the maps $f+g$, $f-g$ and hf are continuous on A, and $\dfrac{f}{h}$ is continuous on $A \backslash \{x \in A \,|\, h(x) = 0\}$.*

Definition 4.22. Let (M,d), (N,ρ) and (P,δ) be metric spaces, A be a subset of M, B be a subset of N, and $f : A \to N$, $g : B \to P$ be maps such that $f(A) \subseteq B$. The composite function $g \circ f : A \to P$ is the map defined by

$$(g \circ f)(x) = g\big(f(x)\big) \qquad \forall\, x \in A\,.$$

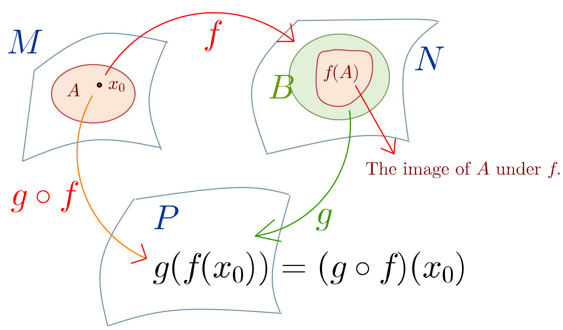


Figure 4.1: The composition of functions

Theorem 4.23. *Let (M,d), (N,ρ) and (P,δ) be metric spaces, A be a subset of M, B be a subset of N, and $f : A \to N$, $g : B \to P$ be maps such that $f(A) \subseteq B$. Suppose that $a \in A \cap A'$ and $\lim\limits_{x \to a} f(x) = b$, and g is continuous at b. Then $\lim\limits_{x \to a} (g \circ f)(x) = g(b)$.*

Proof. Let $\{x_n\}_{n=1}^\infty \subseteq A \backslash \{a\}$ be a convergent sequence with limit a. Then $\{f(x_n\}_{n=1}^\infty$ is a convergent sequence with limit b; thus the continuity of g at b implies that $\big\{g(f(x_n))\big\}_{n=1}^\infty$ converges to $g(b)$. $\square$

Corollary 4.24. *Let (M,d), (N,ρ) and (P,δ) be metric spaces, A be a subset of M, B be a subset of N, and $f : A \to N$, $g : B \to P$ be maps such that $f(A) \subseteq B$.*

1. *If f is continuous at a and g is continuous at $f(a)$, then $g \circ f : A \to P$ is continuous at a.*

2. *If f is continuous on A and g is continuous on B, then $g \circ f : A \to P$ is continuous on A.*

Alternative Proof of 2 *in Corollary* 4.24. Let $\mathcal{W}$ be an open set in (P, r). By Theorem 4.14, there exists $\mathcal{V}$ open in (N, ρ) such that $g^{-1}(\mathcal{W}) = \mathcal{V} \cap B$. Since $\mathcal{V}$ is open in (N, ρ), by Theorem 4.14 again there exists $\mathcal{U}$ open in (M, d) such that $f^{-1}(\mathcal{V}) = \mathcal{U} \cap A$. Then

$$(g \circ f)^{-1}(\mathcal{W}) = f^{-1}\big(g^{-1}(\mathcal{W})\big) = f^{-1}(\mathcal{V} \cap B) = f^{-1}(\mathcal{V}) \cap f^{-1}(B) = \mathcal{U} \cap A \cap f^{-1}(B),$$

while the fact that $f(A) \subseteq B$ further implies that

$$(g \circ f)^{-1}(\mathcal{W}) = \mathcal{U} \cap A.$$

Therefore, by Theorem 4.14 we find that $(g \circ f)$ is continuous on A. □

4.3 Images under Continuous Maps

4.3.1 Image of compact sets

Theorem 4.25. *Let (M, d) and (N, ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a continuous map.*

1. *If $K \subseteq A$ is compact, then $f(K)$ is compact in (N, ρ).*

2. *Moreover, if $(N, \rho) = (\mathbb{R}, |\cdot|)$, then there exist $x_0, x_1 \in K$ such that*

$$f(x_0) = \inf f(K) = \inf \big\{ f(x) \,\big|\, x \in K \big\} \quad \textit{and} \quad f(x_1) = \sup f(K) = \sup \big\{ f(x) \,\big|\, x \in K \big\}.$$

Proof. 1. Let $\{y_n\}_{n=1}^{\infty}$ be a sequence in $f(K)$. Then there exists $\{x_n\}_{n=1}^{\infty} \subseteq K$ such that $y_n = f(x_n)$. Since K is sequentially compact, there exists a convergent subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ with limit $x \in K$. Let $y = f(x) \in f(K)$. By the continuity of f,

$$\lim_{k \to \infty} \rho(y_{n_k}, y) = \lim_{k \to \infty} \rho\big(f(x_{n_k}), f(x)\big) = 0$$

which implies that the sequence $\{y_{n_k}\}_{k=1}^{\infty}$ converges to $y \in f(K)$. Therefore, $f(K)$ is sequentially compact.

2. By 1, $f(K)$ is sequentially compact. Corollary 3.45 then implies that $\inf f(K) \in f(K)$ and $\sup f(K) \in f(K)$. $\square$

Alternative Proof of Part 1. Let $\{V_\alpha\}_{\alpha\in I}$ be an open cover of $f(K)$. Since V_α is open, by Theorem 4.14 there exists U_α open in (M,d) such that $f^{-1}(V_\alpha) = U_\alpha \cap A$. Since $f(K) \subseteq \bigcup_{\alpha\in I} V_\alpha$,

$$K \subseteq f^{-1}(f(K)) \subseteq \bigcup_{\alpha\in I} f^{-1}(V_\alpha) = A \cap \bigcup_{\alpha\in I} U_\alpha$$

which implies that $\{U_\alpha\}_{\alpha\in I}$ is an open cover of K. Therefore,

$$\exists\, J \subseteq I, \#J < \infty \ni K \subseteq A \cap \bigcup_{\alpha\in J} U_\alpha = \bigcup_{\alpha\in J} f^{-1}(V_\alpha);$$

thus $f(K) \subseteq \bigcup_{\alpha\in J} f(f^{-1}(V_\alpha)) \subseteq \bigcup_{\alpha\in J} V_\alpha$. $\square$

Corollary 4.26 (Extreme Value Theorem)**.** *Let $f : [a,b] \to \mathbb{R}$ be continuous. Then f attains its maximum and minimum in $[a,b]$; that is, there are $x_0 \in [a,b]$ and $x_1 \in [a,b]$ such that*

$$f(x_0) = \inf\{f(x) \,|\, x \in [a,b]\} \quad \text{and} \quad f(x_1) = \sup\{f(x) \,|\, x \in [a,b]\}. \tag{4.3.1}$$

Proof. The Heine-Borel Theorem shows that $[a,b]$ is a compact set in $\mathbb{R}$; thus Theorem 4.25 implies that $f([a,b])$ must be compact in $\mathbb{R}$. By Corollary 3.45,

$$\inf f([a,b]) \in f([a,b]) \quad \text{and} \quad \sup f([a,b]) \in f([a,b])$$

that further imply (4.3.1). $\square$

Remark 4.27. If f attains its maximum (or minimum) on a set B, we use $\max\{f(x) \,|\, x \in B\}$ $\big($or $\min\{f(x) \,|\, x \in B\}\big)$ to denote $\sup\{f(x) \,|\, x \in B\}$ $\big($or $\inf\{f(x) \,|\, x \in B\}\big)$. Therefore, (4.3.1) can be rewritten as

$$f(x_0) = \min\{f(x) \,|\, x \in [a,b]\} \quad \text{and} \quad f(x_1) = \max\{f(x) \,|\, x \in [a,b]\}.$$

Remark 4.28. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 0$. Then f is continuous. Note that $\{0\} \subseteq \mathbb{R}$ is compact, but $f^{-1}(\{0\}) = \mathbb{R}$ is not compact. In other words, the pre-image of a compact set under a continuous map might not be compact.

Example 4.29. Recall that two norms $\|\cdot\|$ and $|||\cdot|||$ on a vector space $\mathcal{V}$ are called equivalent if there are positive constants c and C such that

$$c\|x\| \leqslant |||x||| \leqslant C\|x\| \qquad \forall\, x \in \mathcal{V}.$$

We note that equivalent norms on a vector space $\mathcal{V}$ induce the same topology; that is, if $\|\cdot\|$ and $|||\cdot|||$ are equivalent norms on $\mathcal{V}$, then U is open in the normed space $(\mathcal{V}, \|\cdot\|)$ if and only if U is open in the normed space $(\mathcal{V}, |||\cdot|||)$. In fact, let U be an open set in $(\mathcal{V}, \|\cdot\|)$. Then for any $x \in U$, there exists $r > 0$ such that

$$B_{\|\cdot\|}(x, r) \equiv \big\{ y \in \mathcal{V} \,\big|\, \|x - y\| < r \big\} \subseteq U,$$

here we use the norm in the subscript to indicate that the distance in this ball is measured by this norm. As in the proof of Theorem 2.39, the ball $B_{|||\cdot|||}(x, cr) \subseteq B_{\|\cdot\|}(x, r)$. Therefore, U is open in $(\mathcal{V}, |||\cdot|||)$. Similarly, if U is open in $(\mathcal{V}, |||\cdot|||)$, then the inequality $|||x||| \leqslant C_2\|x\|$ implies that U is open in $(\mathcal{V}, \|\cdot\|)$.

In fact, for a vector space $\mathcal{V}$ with two equivalent norms $\|\cdot\|$ and $|||\cdot|||$, we have

1. $\{x_k\}_{k=1}^{\infty}$ converges in $(\mathcal{V}, \|\cdot\|)$ if and only if $\{x_k\}_{k=1}^{\infty}$ converges in $(\mathcal{V}, |||\cdot|||)$.
2. F is (totally) bounded in $(\mathcal{V}, \|\cdot\|)$ if and only if F is a (totally) bounded subset in $(\mathcal{V}, |||\cdot|||)$.
3. F is closed in $(\mathcal{V}, \|\cdot\|)$ if and only if F is closed in $(\mathcal{V}, |||\cdot|||)$.
4. U is open in $(\mathcal{V}, \|\cdot\|)$ if and only if U is open in $(\mathcal{V}, |||\cdot|||)$.
5. K is compact in $(\mathcal{V}, \|\cdot\|)$ if and only if K is compact in $(\mathcal{V}, |||\cdot|||)$.

In the following, we prove the following

Claim: Any two norms on a **finite dimensional** vector space $\mathcal{V}$ over field $\mathbb{R}$ (or $\mathbb{C}$) are equivalent.

Proof of claim: Let $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ be a basis of $\mathcal{V}$. Then each $\boldsymbol{x} \in \mathcal{V}$ can be uniquely expressed as $\boldsymbol{x} = \sum\limits_{i=1}^{n} x^{(i)}\mathbf{e}_i$ for some $x^{(i)} \in \mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Define the norm

$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x^{(i)}|^2}$$

as given in Example 2.28. It suffices to show that any norm $\|\cdot\|$ on $\mathcal{V}$ is equivalent to $\|\cdot\|_2$. In fact, if $C_1\|\boldsymbol{x}\| \leqslant \|\boldsymbol{x}\|_2 \leqslant C_2\|\boldsymbol{x}\|$ and $C_3|||\boldsymbol{x}||| \leqslant \|\boldsymbol{x}\|_2 \leqslant C_4|||\boldsymbol{x}|||$ for all $\boldsymbol{x} \in \mathcal{V}$, then

$$\frac{C_1}{C_4}\|\boldsymbol{x}\| \leqslant |||\boldsymbol{x}||| \leqslant \frac{C_2}{C_3}\|\boldsymbol{x}\| \qquad \forall\, \boldsymbol{x} \in \mathcal{V}.$$

Before proceeding, we first recall that a subset K is (sequentially) compact in $(\mathcal{V}, \|\cdot\|_2)$ if and only if K is closed and bounded (see Remark 3.43). By the triangle inequality and the Cauchy-Schwarz inequality,

$$\|\boldsymbol{x}\| \leqslant \sum_{i=1}^{n} |x^{(i)}|\|\mathbf{e}_i\| \leqslant \|\boldsymbol{x}\|_2 \sqrt{\sum_{i=1}^{n} \|\mathbf{e}_i\|^2}\,; \tag{4.3.2}$$

thus letting $C = \sqrt{\sum\limits_{i=1}^{n} \|\mathbf{e}_i\|^2}$ we have $\|\boldsymbol{x}\| \leqslant C\|\boldsymbol{x}\|_2$.

On the other hand, define $f : \mathcal{V} \to \mathbb{R}$ by

$$f(\boldsymbol{x}) = \|\boldsymbol{x}\| = \Big\|\sum_{i=1}^{n} x^{(i)}\mathbf{e}_i\Big\|.$$

Because of (4.3.2), f is continuous on $(\mathcal{V}, \|\cdot\|_2)$. In fact, for $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$,

$$\big|f(\boldsymbol{x}) - f(\boldsymbol{y})\big| = \big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| \leqslant C\|\boldsymbol{x} - \boldsymbol{y}\|_2$$

which guarantees the continuity of f on $(\mathcal{V}, \|\cdot\|_2)$. Let $K = \big\{\boldsymbol{x} \in \mathcal{V} \,\big|\, \|\boldsymbol{x}\|_2 = 1\big\}$. Then K is sequentially compact in $(\mathcal{V}, \|\cdot\|_2)$ since K is closed and bounded in $(\mathcal{V}, \|\cdot\|_2)$; thus by Theorem 4.25 f attains its minimum on K at some point $\boldsymbol{a} \in K$. Moreover, $f(\boldsymbol{a}) > 0$ (since if $f(\boldsymbol{a}) = 0$, $\boldsymbol{a} = \mathbf{0} \notin K$). Then for all $\boldsymbol{x} \in \mathcal{V}\backslash\{\mathbf{0}\}$, $\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2} \in K$; thus

$$f\Big(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}\Big) \geqslant f(\boldsymbol{a}) \qquad \forall\, \boldsymbol{x} \in \mathcal{V}\backslash\{\mathbf{0}\}.$$

The inequality above further implies that $f(\boldsymbol{a})\|\boldsymbol{x}\|_2 \leqslant f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ for all $\boldsymbol{x} \in \mathcal{V}$; thus letting $c = f(\boldsymbol{a})$ we have $c\|\boldsymbol{x}\|_2 \leqslant \|\boldsymbol{x}\|$.

Having established that every two norms on a finite dimensional vector space over $\mathbb{R}$ (or $\mathbb{C}$), for a finite dimensional normed vector space $(\mathcal{V}, \|\cdot\|)$ over $\mathbb{R}$ (or $\mathbb{C}$) we have the following results:

1. A subset K of $\mathcal{V}$ is compact if and only if K is closed and bounded (because of Remark 3.43).

2. Every bounded sequence in $(\mathcal{V}, \|\cdot\|)$ has a convergent subsequence (for a given bounded sequence, consider the bounded and closed ball $B[0, R]$ for $R \gg 1$ and make use of the sequentially compactness of $B[0, R]$).

3. $(\mathcal{V}, \|\cdot\|)$ is a Banach space; that is, $(\mathcal{V}, \|\cdot\|)$ is complete (every Cauchy sequence is bounded; thus possessing a convergent subsequence so that the convergence of the Cauchy sequence is guaranteed by Proposition 2.58).

Example 4.30. The determinant function $f : \mathcal{M}_{n\times n} \to \mathbb{R}$ defined by $f(A) = \det(A)$ is continuous on $(\mathcal{M}_{n\times n}, \|\cdot\|)$ for any norm $\|\cdot\|$ (thus the set $\mathrm{GL}(n)$ defined in Example 4.16 is open). To see this, we note that $\mathcal{M}_{n\times n}$ is finite dimensional vector space over $\mathbb{R}$; thus the norm $\|\cdot\|$ is equivalent to the norm

$$|||[a_{ij}]||| = \sum_{i,j=1}^{n} |a_{ij}|.$$

Clearly f is continuous on $(\mathcal{M}_{n\times n}, |||\cdot|||)$ since $f(A)$ is the sum of product of entries of A and $|||B - A||| \to 0$ if and only if $b_{ij} \to a_{ij}$ for all $1 \leqslant i, j \leqslant n$. Since $|||B - A||| \to 0$ if and only if $\|B - A\| \to 0$, we conclude that f is continuous on $(\mathcal{M}_{n\times n}, \|\cdot\|)$.

Corollary 4.31. *Let (M, d) be a metric space, K be a compact subset of M, and $f : K \to \mathbb{R}$ be continuous. Then the set*

$$\big\{x \in K \,\big|\, f(x) \text{ is the maximum of } f \text{ on } K\big\}$$

is a non-empty compact set.

Proof. Note that $f(K)$ is compact in $(\mathbb{R}, |\cdot|)$; hence $f(K)$ is closed and bounded so that $\mathrm{M} = \sup f(K)$ exists and $\mathrm{M} \in f(K)$. Then the set defined above is $f^{-1}(\{\mathrm{M}\})$. Moreover,

1. $f^{-1}(\{\mathrm{M}\})$ is non-empty by Theorem 4.25;

2. $f^{-1}(\{\mathrm{M}\})$ is a subset of K; thus By Proposition 3.51 implies that $f^{-1}(\{\mathrm{M}\})$ is totally bounded;

3. By Theorem 4.14, $f^{-1}(\{\mathrm{M}\})$ is closed since $\{\mathrm{M}\}$ is a closed set in $(\mathbb{R}, |\cdot|)$; thus Theorem 3.27 implies that $f^{-1}(\{\mathrm{M}\})$ is complete.

Therefore, Theorem 3.53 shows that $f^{-1}(\{\mathrm{M}\})$ is compact. ∎

4.3.2 Image of connected sets

Theorem 4.32. *Let (M, d) and (N, ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a continuous map. If $C \subseteq A$ is connected, then $f(C)$ is connected in (N, ρ).*

Proof. Suppose that there are two open sets V_1 and V_2 in (N, ρ) such that

(a) $f(C) \cap V_1 \cap V_2 = \varnothing$; (b) $f(C) \cap V_1 \neq \varnothing$; (c) $f(C) \cap V_2 \neq \varnothing$; (d) $f(C) \subseteq V_1 \cup V_2$.

By Theorem 4.14, there are U_1 and U_2 open in (M, d) such that $f^{-1}(V_1) = U_1 \cap A$ and $f^{-1}(V_2) = U_2 \cap A$. By (d),

$$C \subseteq f^{-1}(f(C)) \subseteq f^{-1}(V_1) \cup f^{-1}(V_2) = (U_1 \cup U_2) \cap A \subseteq U_1 \cup U_2 .$$

Moreover, by (a) we find that

$$\begin{aligned} C \cap U_1 \cap U_2 = C \cap (U_1 \cap A) \cap (U_2 \cap A) &= C \cap f^{-1}(V_1) \cap f^{-1}(V_2) \\ &\subseteq f^{-1}(f(C) \cap V_1 \cap V_2) = \varnothing \end{aligned}$$

which implies $C \cap U_1 \cap U_2 = \varnothing$. Finally, (b) implies that for some $x \in C$, $f(x) \in V_1$. Therefore, $x \in f^{-1}(V_1) = U_1 \cap A$ which shows that $x \in U_1$; thus $C \cap U_1 \neq \varnothing$. Similarly, $C \cap U_2 \neq \varnothing$. Therefore, C is disconnected which is a contradiction. □

Corollary 4.33 (Intermediate Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. If $f(a) \neq f(b)$, then for all d in between $f(a)$ and $f(b)$, there exists $c \in (a, b)$ such that $f(c) = d$.*

Proof. The closed interval $[a, b]$ is connected by Theorem 3.68, so Theorem 4.32 implies that $f([a, b])$ must be connected in $\mathbb{R}$. By Theorem 3.68 again, if d is in between $f(a)$ and $f(b)$, then d belongs to $f([a, b])$. Therefore, for some $c \in (a, b)$ we have $f(c) = d$. □

Remark 4.34. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$. Then f is continuous. Note that $C = \{1\}$ is connected, but $f^{-1}(C) = \{1, -1\}$ is not connected. In other words, the pre-image of a connected set under a continuous map might not be connected.

Example 4.35. Let $f : [0, 1] \to [0, 1]$ be continuous. Then there exists $x_0 \in [0, 1]$ such that $f(x_0) = x_0$.

Proof. Let $g(x) = x - f(x)$.

Case 1: $g(0) = 0$ or $g(1) = 0$. Then $x_0 = 0$ or $x_0 = 1$ satisfies $f(x_0) = x_0$.

Case 2: $g(0) \neq 0$ and $g(1) \neq 0$. Then $g(0) < 0$ and $g(1) > 0$; thus by the continuity of $g : [0,1] \to \mathbb{R}$, there exists $x_0 \in [0,1]$ such that $g(x_0) = 0$ which implies the existence of $x_0 \in (0,1)$ satisfying $f(x_0) = x_0$. □

Remark 4.36. Such an x_0 in Example 4.35 is called a ***fixed-point*** of f.

4.4 Uniform Continuity（均勻連續）

Definition 4.37. Let (M,d) and (N,ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a map. For a set $B \subseteq A$, f is said to be ***uniformly continuous on*** B if for any two sequences $\{x_n\}_{n=1}^{\infty}, \{y_n\}_{n=1}^{\infty} \subseteq B$ with the property that $\lim\limits_{n\to\infty} d(x_n, y_n) = 0$, one has $\lim\limits_{n\to\infty} \rho\big(f(x_n), f(y_n)\big) = 0$. In logic notation, f is uniformly continuous on B if

$$\big(\forall \{x_n\}_{n=1}^{\infty}, \{y_n\}_{n=1}^{\infty} \subseteq B\big)\Big(\lim_{n\to\infty} d(x_n, y_n) = 0 \Rightarrow \lim_{n\to\infty} \rho\big(f(x_n), f(y_n)\big) = 0\Big).$$

Proposition 4.38. *Let (M,d) and (N,ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a map. If f is uniformly continuous on A, then f is continuous on A.*

Proof. Let $a \in A \cap A'$, and $\{x_k\}_{k=1}^{\infty} \subseteq A$ be a sequence such that $x_k \to a$ as $k \to \infty$. Let $\{y_k\}_{k=1}^{\infty}$ be a constant sequence with value a; that is, $y_k = a$ for all $k \in \mathbb{N}$. Then $\{y_k\}_{k=1}^{\infty} \subseteq A$ and $d(x_k, y_k) \to 0$ as $k \to \infty$. By the uniform continuity of f on A,

$$\lim_{k\to\infty} \rho\big(f(x_k), f(a)\big) = \lim_{k\to\infty} \rho\big(f(x_k), f(y_k)\big) = 0$$

which implies that f is continuous on a. □

Example 4.39. Let $f : [0,1] \to \mathbb{R}$ be the Dirichlet function; that is,

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q}, \\ 1 & \text{if } x \in \mathbb{Q}^{\complement}. \end{cases}$$

and $B = \mathbb{Q} \cap [0,1]$. Then f is continuous **nowhere** in $[0,1]$, but f is uniformly continuous on B. However, the proposition above guarantees that if f is uniformly continuous on $[0,1]$, then f must be continuous on $[0,1]$ (Check why the proof of Proposition 4.38 does not go through if B is a proper subset of $[0,1]$).

Example 4.40. The function $f(x) = |x|$ is uniformly continuous on $\mathbb{R}$. In fact, by the triangle inequality,

$$\big|f(x) - f(y)\big| = \big||x| - |y|\big| \leqslant |x - y|\,;$$

thus if $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ are sequences in $\mathbb{R}$ and $\lim\limits_{n\to\infty} |x_n - y_n| = 0$, by the Sandwich lemma we must have $\lim\limits_{n\to\infty} \big|f(x_n) - f(y_n)\big| = 0$.

Example 4.41. The function $f : (0, \infty) \to \mathbb{R}$ defined by $f(x) = \dfrac{1}{x}$ is uniformly continuous on $[a, \infty)$ for all $a > 0$. To see this, let $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ be sequences in $[a, \infty)$ such that $\lim\limits_{n\to\infty} |x_n - y_n| = 0$. Then

$$|f(x_n) - f(y_n)| = \Big|\frac{1}{x_n} - \frac{1}{y_n}\Big| = \frac{|x_n - y_n|}{|x_n y_n|} \leqslant \frac{|x_n - y_n|}{a^2} \to 0 \quad \text{as} \quad n \to \infty$$

which implies that f is uniformly continuous on $[a, \infty)$ if $a > 0$.

However, f is not uniformly continuous on $(0, \infty)$. Let $x_n = \dfrac{1}{n}$ and $y_n = \dfrac{1}{2n}$. Then

$$|x_n - y_n| = \frac{1}{2n} \to 0 \ \text{ as } \ n \to \infty \quad \text{but} \quad |f(x_n) - f(y_n)| = n \geqslant 1\,.$$

Example 4.42. Let $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. Then f is continuous in $\mathbb{R}$ but not uniformly continuous on $\mathbb{R}$. Let $x_n = n$ and $y_n = n + \dfrac{1}{2n}$. Then

$$\big|f(x_n) - f(y_n)\big| = \Big|n^2 - (n + \frac{1}{2n})^2\Big| = \Big|n^2 - n^2 - 1 - \frac{1}{4n^2}\Big| = 1 + \frac{1}{4n^2} \nrightarrow 0 \ \text{ as } \ n \to \infty\,.$$

Example 4.43. The function $f(x) = \sin(x^2)$ is not uniform continuous on $\mathbb{R}$. Define $x_n = 2n\sqrt{\pi} + \dfrac{\sqrt{\pi}}{8n}$ and $y_n = 2n\sqrt{\pi} - \dfrac{\sqrt{\pi}}{8n}$. Then $\lim\limits_{n\to\infty} |x_n - y_n| = 0$ while

$$\big|\sin(x_n^2) - \sin(y_n^2)\big| = \Big|\sin\big(4n^2\pi + \frac{\pi}{2} + \frac{\pi}{64n^2}\big) - \sin\big(4n^2\pi - \frac{\pi}{2} + \frac{\pi}{64n^2}\big)\Big| = 2\cos\frac{\pi}{64n^2}\,;$$

thus $\lim\limits_{n\to\infty} \big|\sin(x_n^2) - \sin(y_n^2)\big| = 1 \neq 0$.

Example 4.44. The function $f : (0, 1) \to \mathbb{R}$ defined by $f(x) = \sin\dfrac{1}{x}$ is not uniformly continuous on $(0, 1)$.

Let $x_n = \big(2n\pi + \frac{\pi}{2}\big)^{-1}$ and $y_n = \big(2n\pi - \frac{\pi}{2}\big)^{-1}$. Then

$$\Big|\sin\frac{1}{x_n} - \sin\frac{1}{y_n}\Big| = 2\,,$$

while $|x_n - y_n| = \dfrac{\pi}{4n^2\pi^2 - \frac{\pi^2}{4}} = \dfrac{1}{(4n^2 - \frac{1}{4})\pi} \to 0$ as $n \to \infty$.

Theorem 4.45. *Let (M,d) and (N,ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a map. For a set $B \subseteq A$, f is uniformly continuous on B if and only if*

$$\forall\, \varepsilon > 0, \exists\, \delta > 0 \ni \rho\big(f(x), f(y)\big) < \varepsilon \quad \textit{whenever} \quad d(x,y) < \delta \textit{ and } x, y \in B\,.$$

Proof. "$\Leftarrow$" Let $\{x_n\}_{n=1}^\infty$, $\{y_n\}_{n=1}^\infty$ be sequences in B such that $\lim\limits_{n\to\infty} d(x_n, y_n) = 0$, and $\varepsilon > 0$ be given. By assumption, there exists $\delta > 0$ such that

$$\rho\big(f(x), f(y)\big) < \varepsilon \quad \text{whenever} \quad d(x,y) < \delta \text{ and } x, y \in B\,.$$

Since $\lim\limits_{n\to\infty} d(x_n, y_n) = 0$, there exists $N > 0$ such that

$$d(x_n, y_n) < \delta \quad \text{whenever} \quad n \geqslant N\,;$$

thus

$$\rho\big(f(x_n), f(y_n)\big) < \varepsilon \quad \text{whenever} \quad n \geqslant N\,.$$

"$\Rightarrow$" Suppose the contrary that there exists $\varepsilon > 0$ such that for all $\delta = \dfrac{1}{n} > 0$, there exist two points x_n and $y_n \in B$ such that

$$d(x_n, y_n) < \frac{1}{n} \quad \text{but} \quad \rho\big(f(x_n), f(y_n)\big) \geqslant \varepsilon\,.$$

These points form two sequences $\{x_n\}_{n=1}^\infty$, $\{y_n\}_{n=1}^\infty$ in B such that $\lim\limits_{n\to\infty} d(x_n, y_n) = 0$, while the limit of $\rho\big(f(x_n), f(y_n)\big)$, if exists, does not converges to zero as $n \to \infty$. As a consequence, f is not uniformly continuous on B, a contradiction. $\square$

Remark 4.46. The theorem above provides another way (the blue color part) of defining the uniform continuity of a function over a subset of its domain. Moreover, according to this alternative definition, $f : A \to N$ is uniformly continuous on $B \subseteq A$ if

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)\Big(\operatorname{diam}\Big(f\Big(B_M\big(b, \frac{\delta}{2}\big) \cap B\Big)\Big) < \varepsilon\Big);$$

that is, the diameter of the image, under f, of subsets of B whose diameter is not greater than δ is not greater than ε（在 B 中直徑不超過 δ 的子集合被函數 f 映過去之後，在對應域中的直徑不會超過 ε）. The statement above is the same as

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)(\forall\, b \in M)(\exists\, c \in M)\Big(f\Big(B_M\big(b, \frac{\delta}{2}\big) \cap B\Big) \subseteq B_N\big(c, \frac{\varepsilon}{2}\big)\Big)\,.$$

Remark 4.47. For a given function f, let $\delta(f,c,\varepsilon)$ denote the supremum of all $\delta(c,\varepsilon)$ mentioned in Remark 4.6. Then the uniform continuity of a function $f : A \to N$ is equivalent to that

$$\delta_f(\varepsilon) \equiv \inf_{c\in A} \delta(f,c,\varepsilon) > 0 \qquad \forall\, \varepsilon > 0\,.$$

Remark 4.48. Let (M,d) and (N,ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a map. For a set $B \subseteq A$, the following four statements are equivalent:

(1) f is **not** uniformly continuous on B.

(2) $\exists\, \{x_n\}_{n=1}^\infty, \{y_n\}_{n=1}^\infty \subseteq B \ni \lim\limits_{n\to\infty} d(x_n,y_n) = 0$ and $\limsup\limits_{n\to\infty} \rho\big(f(x_n), f(y_n)\big) > 0$.

(3) $\exists\, \{x_n\}_{n=1}^\infty, \{y_n\}_{n=1}^\infty \subseteq B \ni \lim\limits_{n\to\infty} d(x_n,y_n) = 0$ and $\lim\limits_{n\to\infty} \rho\big(f(x_n), f(y_n)\big) > 0$.

(4) $\exists\, \varepsilon > 0 \ni \forall\, n > 0, \exists\, x_n, y_n \in B$ and $d(x_n,y_n) < \dfrac{1}{n} \ni \rho\big(f(x_n), f(y_n)\big) \geqslant \varepsilon$.

Theorem 4.49. *Let (M,d) and (N,ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a map. If K is a compact subset of A and f is continuous on K, then f is uniformly continuous on K.*

Proof. Assume the contrary that f is not uniformly continuous on K. Then ((3) of Remark 4.48 implies that) there are sequences $\{x_n\}_{n=1}^\infty$ and $\{y_n\}_{n=1}^\infty$ in K such that

$$\lim_{n\to\infty} d(x_n,y_n) = 0 \quad \text{but} \quad \lim_{n\to\infty} \rho\big(f(x_n), f(y_n)\big) > 0\,.$$

Since K is (sequentially) compact, there exist convergent subsequences $\{x_{n_k}\}_{k=1}^\infty$ and $\{y_{n_k}\}_{k=1}^\infty$ with limits $x, y \in K$. On the other hand, $\lim\limits_{n\to\infty} d(x_n,y_n) = 0$, we must have $x = y$; thus by the continuity of f (on K),

$$0 = \rho\big(f(x), f(x)\big) = \lim_{k\to\infty} \rho\big(f(x_{n_k}), f(y_{n_k})\big) = \lim_{n\to\infty} \rho\big(f(x_n), f(y_n)\big) > 0\,,$$

a contradiction. $\square$

Alternative proof. Let $\varepsilon > 0$ be given. Since f is continuous on K,

$$\forall\, a \in K, \exists\, \delta = \delta(a) > 0 \ni \rho\big(f(x), f(a)\big) < \frac{\varepsilon}{2} \text{ whenever } x \in B(a,\delta) \cap A\,.$$

Then $\Big\{B\big(a, \dfrac{\delta(a)}{2}\big)\Big\}_{a\in K}$ is an open cover of K; thus the compactness of K implies that

$$\exists\, \{a_1, \cdots, a_N\} \subseteq K \ni K \subseteq \bigcup_{i=1}^N B\big(a_i, \frac{\delta_i}{2}\big)\,,$$

where $\delta_i = \delta(a_i)$. Let $\delta = \dfrac{1}{2}\min\{\delta_1, \cdots, \delta_N\}$. Then $\delta > 0$, and if $x_1, x_2 \in K$ and $d(x_1, x_2) < \delta$, there must be $j = 1, \cdots, N$ such that $x_1, x_2 \in B(a_j, \delta_j)$. In fact, since $x_1 \in B\big(a_j, \dfrac{\delta_j}{2}\big)$ for some $j = 1, \cdots, N$, then

$$d(x_2, a_j) \leqslant d(x_1, x_2) + d(x_1, a_j) < \delta + \frac{\delta_j}{2} \leqslant \delta_j\,.$$

Therefore, $x_1, x_2 \in B(a_j, \delta_j) \cap A$ for some $j = 1, \cdots, N$; thus

$$\rho\big(f(x_1), f(x_2)\big) \leqslant \rho\big(f(x_1), f(a_j)\big) + \rho\big(f(x_2), f(a_j)\big) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,. \qquad \square$$

Lemma 4.50. *Let (M, d) and (N, ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be uniformly continuous. If $\{x_k\}_{k=1}^\infty \subseteq A$ is a Cauchy sequence, so is $\big\{f(x_k)\big\}_{k=1}^\infty$.*

Proof. Let $\{x_k\}_{k=1}^\infty$ be a Cauchy sequence in (M, d), and $\varepsilon > 0$ be given. Since $f : A \to N$ is uniformly continuous,

$$\exists\, \delta > 0 \ni \rho\big(f(x), f(y)\big) < \varepsilon \quad \text{whenever} \quad d(x, y) < \delta \text{ and } x, y \in A\,.$$

For this particular δ, there exists $N > 0$ such that $d(x_k, x_\ell) < \delta$ whenever $k, \ell \geqslant N$. Therefore,

$$\rho(f(x_k), f(x_\ell)\big) < \varepsilon \quad \text{whenever} \quad k, \ell \geqslant N\,. \qquad \square$$

Corollary 4.51. *Let (M, d) and (N, ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be uniformly continuous. If N is complete, then f has a unique extension to a continuous function on $\bar{A}$; that is, there exists $g : \bar{A} \to N$ such that*

1. *g is uniformly continuous on $\bar{A}$;*
2. *$g(x) = f(x)$ for all $x \in A$;*
3. *if $h : \bar{A} \to N$ is a continuous map satisfying $h(x) = f(x)$ for all $x \in A$, then $h = g$.*

Proof. Let $x \in \bar{A}\backslash A$. Then there exists $\{x_k\}_{k=1}^\infty \subseteq A$ such that $x_k \to x$ as $k \to \infty$. Since $\{x_k\}_{k=1}^\infty$ is Cauchy, by Lemma 4.50 $\big\{f(x_k)\big\}_{k=1}^\infty$ is a Cauchy sequence in (N, ρ); thus is convergent. Moreover, if $\{z_k\}_{k=1}^\infty \subseteq A$ is another sequence converging to x, we must have $d(x_k, z_k) \to 0$ as $k \to \infty$; thus $\rho(f(x_k), f(z_k)) \to 0$ as $k \to \infty$, so the limit of $\big\{f(x_k)\big\}_{k=1}^\infty$ and $\big\{f(z_k)\big\}_{k=1}^\infty$ must be the same.

Define $g : \bar{A} \to N$ by

$$g(x) = \begin{cases} f(x) & \text{if } x \in A\,, \\ \lim\limits_{k\to\infty} f(x_k) & \text{if } x \in \bar{A}\backslash A, \text{ and } \{x_k\}_{k=1}^\infty \subseteq A \text{ converging to } x \text{ as } k \to \infty\,. \end{cases}$$

Then the argument above shows that g is well-defined, and (2) holds.

Let $\varepsilon > 0$ be given. Since $f : A \to N$ is uniformly continuous,

$$\exists\, \delta > 0 \ni \rho\big(f(x), f(y)\big) < \frac{\varepsilon}{3} \text{ whenever } d(x,y) < 2\delta \text{ and } x, y \in A.$$

Suppose that $x, y \in \bar{A}$ such that $d(x,y) < \delta$. Let $\{x_k\}_{k=1}^\infty, \{y_k\}_{k=1}^\infty \subseteq A$ be sequences converging to x and y, respectively. Then there exists $N > 0$ such that

$$d(x_k, x) < \frac{\delta}{2}, d(y_k, y) < \frac{\delta}{2} \text{ and } \rho\big(f(x_k), g(x)\big) < \frac{\varepsilon}{3}, \rho\big(f(y_k), g(y)\big) < \frac{\varepsilon}{3} \quad \forall\, k \geqslant N.$$

In particular, due to the triangle inequality,

$$d(x_N, y_N) \leqslant d(x_N, x) + d(x,y) + d(y, y_N) < \frac{\delta}{2} + \delta + \frac{\delta}{2} = 2\delta;$$

thus $\rho\big(f(x_N), f(y_N)\big) < \frac{\varepsilon}{3}$. As a consequence,

$$\rho\big(g(x), g(y)\big) \leqslant \rho\big(g(x), f(x_N)\big) + \rho\big(f(x_N), f(y_N)\big) + \rho\big(f(y_N), f(y)\big) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

which establishes (1).

Finally, suppose that $h : \bar{A} \to N$ is a continuous map satisfying $h = f$ on A, and $a \in A$. Let $\{x_k\}_{k=1}^\infty$ be a sequence in A with limit a. By Proposition 4.38, g is continuous on $\bar{A}$; thus Proposition 4.9 implies that

$$g(a) = \lim_{k\to\infty} g(x_k) = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} h(x_k) = h(a),$$

so (3) is also concluded. $\square$

4.5 Exercises

§ 4.1 Continuity

Started from this section, for all $n \in \mathbb{N}$ $\mathbb{R}^n$ always denotes the normed space $(\mathbb{R}^n, \|\cdot\|_2)$.

Problem 4.1. Use whatever methods you know to find the following limits:

1. $\displaystyle\lim_{x\to 0^+} (1 + \sin 2x)^{\frac{1}{x}}$;
2. $\displaystyle\lim_{x\to -\infty} \big(\sqrt{1+x+x^2} - \sqrt{1-x+x^2}\big)$;
3. $\displaystyle\lim_{x\to 1} (2-x)^{\sec\frac{\pi x}{2}}$;
4. $\displaystyle\lim_{x\to\infty} x\Big(\frac{\pi}{2} - \sin^{-1}\frac{x}{\sqrt{x^2+1}}\Big)$;
5. $\displaystyle\lim_{x\to\infty} x\bigg(e^{-1} - \Big(\frac{x}{x+1}\Big)^x\bigg)$;
6. $\displaystyle\lim_{x\to\infty} \Big(\frac{a^x - 1}{x(a-1)}\Big)^{\frac{1}{x}}$, where $a > 0$ and $a \neq 1$.

Problem 4.2. Complete the following.

1. Find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that
$$\lim_{x\to 0}\lim_{y\to 0} f(x,y) \quad \text{and} \quad \lim_{y\to 0}\lim_{x\to 0} f(x,y)$$
exist but are not equal.

2. Find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that the two limits above exist and are equal but f is not continuous.

3. Find a function $f : \mathbb{R}^2 \to \mathbb{R}$ that is continuous on every line through the origin but is not continuous.

Problem 4.3. Complete the following.

1. Show that the projection map $f : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x \end{array}$ is continuous.

2. Show that if $U \subseteq \mathbb{R}$ is open, then $A = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, x \in U\right\}$ is open.

3. Give an example of a continuous function $f : \mathbb{R} \to \mathbb{R}$ and an open set $U \subseteq \mathbb{R}$ such that $f(U)$ is not open.

Problem 4.4. Show that $f : A \to \mathbb{R}^m$, where $A \subseteq \mathbb{R}^n$, is continuous if and only if for every $B \subseteq A$,
$$f(\mathrm{cl}(B) \cap A) \subseteq \mathrm{cl}(f(B))\,.$$

Problem 4.5. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ satisfy $T(x+y) = T(x) + T(y)$ for all $x, y \in \mathbb{R}^n$.

1. Show that $T(rx) = rT(x)$ for all $r \in \mathbb{Q}$ and $x \in \mathbb{R}^n$.

2. Suppose that T is continuous on $\mathbb{R}^n$. Show that T is linear; that is, $T(cx+y) = cT(x) + T(y)$ for all $c \in \mathbb{R}$, $x, y \in \mathbb{R}^n$.

3. Suppose that T is continuous at some point x_0 in $\mathbb{R}^n$. Show that T is continuous on $\mathbb{R}^n$.

4. Suppose that T is bounded on some open subset of $\mathbb{R}^n$. Show that T is continuous on $\mathbb{R}^n$.

5. Suppose that T is bounded from above (or below) on some open subset of $\mathbb{R}^n$. Show that T is continuous on $\mathbb{R}^n$.

6. Construct a $T : \mathbb{R} \to \mathbb{R}$ which is discontinuous at every point of $\mathbb{R}$, but $T(x+y) = T(x) + T(y)$ for all $x, y \in \mathbb{R}$.

Problem 4.6. Let C be a convex set in $\mathbb{R}^n$ equipped with inner product $\langle\cdot,\cdot\rangle$ (for the definition of convex sets, see Problem 2.5), and $f : C \to \mathbb{R}$ be a convex function; that is,

$$f(tx+(1-t)y) \leqslant tf(x)+(1-t)f(y) \qquad \forall\, t \in [0,1] \text{ and } x, y \in C\,.$$

Complete the following.

1. Show that f is locally bounded (from above); that is, for every $x_0 \in \operatorname{int}(C)$ there exists $\delta > 0$ such that f is bounded (from above) on $B(x_0, \delta)$.

2. Show that f is locally Lipschitz; that is, for every $x_0 \in \operatorname{int}(C)$ there exists $\delta > 0$ and $M = M_\delta > 0$ such that

$$\big|f(x)-f(y)\big| \leqslant M_\delta \|x-y\| \qquad \forall\, x, y \in B(x_0, \delta)\,.$$

   In particular, a convex function (defined on a subset of $\mathbb{R}^n$) is continuous in the interior of its domain.

3. For $x_0 \in \operatorname{int}(C)$, define the sub-differential

$$\partial f(x_0) = \Big\{ p \in \mathbb{R}^n \,\Big|\, f(x) \geqslant f(x_0) + \langle p, x - x_0\rangle \text{ for all } x \in C \Big\}.$$

   Show that $\partial f(x_0)$ is non-empty, convex and compact.

4. Show that f attains a local minimum at $x_0 \in \operatorname{int}(C)$ if and only if $0 \in \partial f(x_0)$.

Problem 4.7. Let (M, d) be a metric space, $A \subseteq M$, and $f : A \to \mathbb{R}$. For $a \in A'$, define

$$\liminf_{x\to a} f(x) = \lim_{r\to 0^+} \inf\big\{ f(x) \,\big|\, x \in B(a,r) \cap A\backslash\{a\}\big\}\,,$$
$$\limsup_{x\to a} f(x) = \lim_{r\to 0^+} \sup\big\{ f(x) \,\big|\, x \in B(a,r) \cap A\backslash\{a\}\big\}\,.$$

Complete the following.

1. Show that both $\liminf\limits_{x\to a} f(x)$ and $\limsup\limits_{x\to a} f(x)$ exist (which may be $\pm\infty$), and
$$\liminf_{x\to a} f(x) \leqslant \limsup_{x\to a} f(x)\,.$$
Furthermore, there exist sequences $\{x_n\}_{n=1}^\infty$, $\{y_n\}_{n=1}^\infty \subseteq A\backslash\{a\}$ such that $\{x_n\}_{n=1}^\infty$ and $\{y_n\}_{n=1}^\infty$ both converge to a, and
$$\lim_{n\to\infty} f(x_n) = \liminf_{x\to a} f(x) \qquad \text{and} \qquad \lim_{n\to\infty} f(y_n) = \limsup_{x\to a} f(x)\,.$$

2. Let $\{x_n\}_{n=1}^\infty \subseteq A\backslash\{a\}$ be a convergent sequence with limit a. Show that
$$\liminf_{x\to a} f(x) \leqslant \liminf_{n\to\infty} f(x_n) \leqslant \limsup_{n\to\infty} f(y_n) \leqslant \limsup_{x\to a} f(x)\,.$$

3. Show that $\lim\limits_{x\to a} f(x) = \ell$ if and only if
$$\liminf_{x\to a} f(x) = \limsup_{x\to a} f(x) = \ell\,.$$

4. Show that $\liminf\limits_{x\to a} f(x) = \ell \in \mathbb{R}$ if and only if the following two conditions hold:
   (a) for all $\varepsilon > 0$, there exists $\delta > 0$ such that $\ell - \varepsilon < f(x)$ for all $x \in B(a,\delta) \cap A\backslash\{a\}$;
   (b) for all $\varepsilon > 0$ and $\delta > 0$, there exists $x \in B(a,\delta) \cap A\backslash\{a\}$ such that $f(x) < \ell + \varepsilon$.

   Formulate a similar criterion for limsup and for the case that $\ell = \pm\infty$.

5. Compute the liminf and limsup of the following functions at any point of $\mathbb{R}$.
   (a) $f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q}^\complement\,, \\ \dfrac{1}{p} & \text{if } x = \dfrac{q}{p} \text{ with } (p,q) = 1,\ q > 0,\ p \neq 0\,. \end{cases}$
   (b) $f(x) = \begin{cases} x & \text{if } x \in \mathbb{Q}\,, \\ -x & \text{if } x \in \mathbb{Q}^\complement\,. \end{cases}$

Problem 4.8. Let (M,d) be a metric space, and $A \subseteq M$. A function $f : A \to \mathbb{R}$ is called ***lower semi-continuous*** / ***upper semi-continuous*** at $a \in A$ if either $a \in A\backslash A'$ or $\liminf\limits_{x\to a} f(x) \geqslant f(a)$, / $\limsup\limits_{x\to a} f(x) \leqslant f(a)$, and is called lower/upper semi-continuous on A if f is lower/uppser semi-continuous at a for all $a \in A$.

1. Show that $f : A \to \mathbb{R}$ is lower semi-continuous on A if and only if $f^{-1}((-\infty, r])$ is closed relative to A. Also show that $f : A \to \mathbb{R}$ is upper semi-continuous on A if and only if $f^{-1}([r, \infty))$ is closed relative to A.

2. Show that f is lower semi-continuous on A if and only if for all convergent sequences $\{x_n\}_{n=1}^{\infty} \subseteq A$ and $\{s_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ satisfying $f(x_n) \leqslant s_n$ for all $n \in \mathbb{N}$, we have
$$f\big(\lim_{n\to\infty} x_n\big) \leqslant \lim_{n\to\infty} s_n .$$

3. Let $\{f_\alpha\}_{\alpha\in I}$ be a family of lower semi-continuous functions on A. Prove that $f(x) = \sup\limits_{\alpha\in I} f_\alpha(x)$ is lower semi-continuous on A.

4. Let A be a perfect set (that is, A contains no isolated points) and $f : A \to \mathbb{R}$ be given. Define
$$f^*(x) = \limsup_{y\to x} f(y) \qquad \text{and} \qquad f_*(x) = \liminf_{y\to x} f(y) .$$
Show that f^* is upper semi-continuous and f_* is lower semi-continuous, and $f_*(x) \leqslant f(x) \leqslant f^*(x)$ for all $x \in A$. Moreover, if g is a lower semi-continuous function on A such that $g(x) \leqslant f(x)$ for all $x \in A$, then $g \leqslant f_*$.

§ 4.3 Images under Continuous Maps

Problem 4.9. Complete the following.

1. Show that if $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous, and $B \subseteq \mathbb{R}^n$ is bounded, then $f(B)$ is bounded.

2. If $f : \mathbb{R} \to \mathbb{R}$ is continuous and $K \subseteq \mathbb{R}$ is compact, is $f^{-1}(K)$ necessarily compact?

3. If $f : \mathbb{R} \to \mathbb{R}$ is continuous and $C \subseteq \mathbb{R}$ is connected, is $f^{-1}(C)$ necessarily connected?

Problem 4.10. Let (M, d) be a metric space, $K \subseteq M$ be compact, and $f : K \to \mathbb{R}$ be lower semi-continuous (see Problem 4.8 for the definition). Show that f attains its minimum on K.

Problem 4.11. Let (M, d) be a metric space. Show that a subset $A \subseteq M$ is connected if and only if every continuous function defined on A whose range is a subset of $\{0, 1\}$ is constant.

Problem 4.12. Let $\mathcal{D} \subseteq \mathbb{R}^n$ be an open connected set, where $n > 1$. If a, b and c are any three points in $\mathcal{D}$, show that there is a path in G which connects a and b without passing through c. In particular, this shows that $\mathcal{D}$ is path connected and $\mathcal{D}$ is not homeomorphic to any subset of $\mathbb{R}$.

Problem 4.13. Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \|x\|$. Show that f is continuous on $(\mathbb{R}^n, \|\cdot\|_2)$.
Hint: Show that $|f(x) - f(y)| \leqslant C\|x - y\|_2$ for some fixed constant $C > 0$.

Problem 4.14. Let $(\mathcal{V}, \|\cdot\|)$ be a finite dimensional normed vector fields over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\dim(\mathcal{V}) < \infty$. Show that a subset K of $\mathcal{V}$ is compact if and only if K is closed and bounded.
Hint: See Problem 3.21 for the case $\|\cdot\| = \|\cdot\|_2$, and the general case follows from Example 4.29.

Problem 4.15. Let $C \subseteq \mathbb{R}^n$ be a convex set, and $f : C \to \mathbb{R}$ be a convex function. Show that

1. f is locally bounded from above; that is, for all $\boldsymbol{x}_0 \in \text{int}(C)$ there exists $\delta > 0$ such that f is bounded from above on $B(\boldsymbol{x}_0, \delta)$.

2. f is locally Lipschitz; that is, for all $\boldsymbol{x}_0 \in \text{int}(C)$ there exist $\delta > 0$ and $M = M_\delta > 0$ such that
$$\big|f(\boldsymbol{x}) - f(\boldsymbol{y})\big| \leqslant M_\delta \|\boldsymbol{x} - \boldsymbol{y}\| \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in B(\boldsymbol{x}_0, \delta).$$

Remark 4.52. By part 2 of Problem 4.15 we know that a convex function is continuous in the interior of its domain. In particular, under the setting of Problem 4.15, if $\boldsymbol{x} \in \text{int}(C)$, then $(\boldsymbol{x}, f(\boldsymbol{x}))$ is a boundary point of the set
$$X \equiv \Big\{(\boldsymbol{x}, x^{(n)}) \in \mathbb{R}^{n+1} \,\Big|\, \boldsymbol{x} \in C, x^{(n)} \geqslant f(\boldsymbol{x})\Big\}.$$

In Exercise Problem 4.16 through 4.19, we focus on another kind of connected sets, so-called path connected sets. First we introduce path connectedness in the following

Definition 4.53. Let (M, d) be a metric space. A subset $A \subseteq M$ is said to be ***path connected*** if for every $x, y \in A$, there exists a continuous map $\varphi : [0, 1] \to A$ such that $\varphi(0) = x$ and $\varphi(1) = y$.

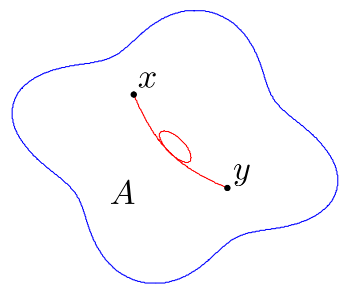


Figure 4.2: Path connected sets

Problem 4.16. Recall that a set A in a vector space $\mathcal{V}$ is called convex if for all $x, y \in A$, the line segment joining x and y lies in A. Show that a convex set in a normed space is path connected.

Problem 4.17. A set S in a vector space $\mathcal{V}$ is called ***star-shaped*** if there exists $p \in S$ such that for any $q \in S$, the line segment joining p and q lies in S. Show that a star-shaped set in a normed space is path connected.

Problem 4.18. Let $A = \left\{\left(x, \sin\frac{1}{x}\right) \,\middle|\, x \in (0,1]\right\} \cup (\{0\} \times [-1,1])$. Show that A is connected in $(\mathbb{R}^2, \|\cdot\|_2)$, but A is not path connected.

Problem 4.19. Let (M, d) be a metric space, and $A \subseteq M$. Show that if A is path connected, then A is connected.
Hint: Apply Theorem 3.68 and prove by contradiction.

Problem 4.20. Let (M, d), (N, ρ) be metric spaces, A be a subset of M, and $f : A \to N$ be a continuous map. Show that if $C \subseteq A$ is path connected, so is $f(C)$.

§ 4.4 Uniform Continuity

Problem 4.21. Check if the following functions on uniformly continuous.

1. $f : (0, \infty) \to \mathbb{R}$ defined by $f(x) = \sin \log x$.
2. $f : (0, 1) \to \mathbb{R}$ defined by $f(x) = x \sin \frac{1}{x}$.
3. $f : (0, \infty) \to \mathbb{R}$ defined by $f(x) = \sqrt{x}$.
4. $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \cos(x^2)$.
5. $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \cos^3 x$.

6. $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x \sin x$.

Problem 4.22. 1. Find all positive numbers a and b such that the function $f(x) = \dfrac{\sin(x^a)}{1 + x^b}$ is uniformly continuous on $[0, \infty)$.

2. Find all positive numbers a and b such that the function $f(x, y) = |x|^a|y|^b$ is uniformly continuous on $\mathbb{R}^2$.

Problem 4.23. Show that $f : [0, 1] \times \mathbb{R} \to \mathbb{R}$ defined by $f(x, y) = \dfrac{\sqrt{x}}{1 + x^2y^2}$ is uniformly continuous on its domain.

Problem 4.24. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be continuous, and $\lim\limits_{|x| \to \infty} f(x) = b$ exists for some $b \in \mathbb{R}^m$. Show that f is uniformly continuous on $\mathbb{R}^n$.

Problem 4.25. Suppose that $f : \mathbb{R}^n \to \mathbb{R}^m$ is uniformly continuous. Show that there exists $a > 0$ and $b > 0$ such that $\|f(x)\|_{\mathbb{R}^m} \leqslant a\|x\|_{\mathbb{R}^n} + b$.

Problem 4.26. Let $f(x) = \dfrac{q(x)}{p(x)}$ be a rational function define on $\mathbb{R}$, where p and q are two polynomials. Show that f is uniformly continuous on $\mathbb{R}$ if and only if the degree of q is not more than the degree of p plus 1.

Problem 4.27. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuous periodic function; that is, there exists $p > 0$ such that $f(x + p) = f(x)$ for all $x \in \mathbb{R}$ (and f is continuous). Show that f is uniformly continuous on $\mathbb{R}$.

Problem 4.28. Let $(a, b) \subseteq \mathbb{R}$ be an open interval, and $f : (a, b) \to \mathbb{R}^m$ be a function. Show that the following three statements are equivalent.

1. f is uniformly continuous on (a, b).

2. f is continuous on (a, b), and both limits $\lim\limits_{x \to a^+} f(x)$ and $\lim\limits_{x \to b^-} f(x)$ exist.

3. For all $\varepsilon > 0$, there exists $N > 0$ such that $\big|f(x) - f(y)\big| < \varepsilon$ whenever $\left|\dfrac{f(x) - f(y)}{x - y}\right| > N$ and $x, y \in (a, b)$, $x \neq y$.

Problem 4.29. Suppose that $f : [a, b] \to \mathbb{R}$ is ***Hölder continuous with exponent*** α; that is, there exist $M > 0$ and $\alpha \in (0, 1]$ such that

$$|f(x) - f(y)| \leqslant M|x - y|^\alpha \qquad \forall\, x, y \in [a, b]\,.$$

Show that f is uniformly continuous on $[a, b]$. Show that $f : [0, \infty) \to \mathbb{R}$ defined by $f(x) = \sqrt{x}$ is Hölder continuous with exponent $\dfrac{1}{2}$.

Problem 4.30. A function $f : A \times B \to \mathbb{R}^m$, where $A \subseteq \mathbb{R}$ and $B \subseteq \mathbb{R}^p$, is said to be separately continuous if for each $x_0 \in A$, the map $g(y) = f(x_0, y)$ is continuous and for $y_0 \in B$, $h(x) = f(x, y_0)$ is continuous. f is said to be continuous on A uniformly with respect to B if

$$\forall\, \varepsilon > 0, \exists\, \delta > 0 \ni \big\|f(x, y) - f(x_0, y)\big\|_2 < \varepsilon \quad \text{whenever} \quad \|x - x_0\|_2 < \delta \text{ and } y \in B\,.$$

Show that if f is separately continuous and is continuous on A uniformly with respect to B, then f is continuous on $A \times B$.

Problem 4.31. Let (M, d) be a metric space, $A \subseteq M$, and $f, g : A \to \mathbb{R}$ be uniformly continuous on A. Show that if f and g are bounded, then fg is uniformly continuous on A. Does the conclusion still hold if f or g is not bounded?

§ 4.5 Chapter Review

Problem 4.32. Two metric spaces (M, d) and (N, ρ) are called ***homeomorphics*** if there exists a continous map $f : M \to N$, called a ***homeomorphism*** between M and N, such that f is one-to-one and onto, and its inverse f^{-1} is also continuous. Homeomorphic metric spaces have the same topological properties. In the following problems, (M, d) and (N, ρ) are two metric spaces.

1. Suppose that M is compact, and $f : M \to N$ is one-to-one and onto. Show that f is a homeomorphism between M and N.

2. Suppose that f is a homeomorphism between M and N. Show that the restriction of f to any subset $A \subseteq M$ establishes a homeomorphism between A and $f(A)$.

3. Determine which of the following pairs of metric spaces is homeomorphic.

   (a) $M = (a, b] \subseteq \mathbb{R}$ and $N = \mathbb{R}$.

   (b) M is an open ball in $\mathbb{R}^n$ and $N = \mathbb{R}^n$.

   (c) $M = \mathbb{R}$ and $N = \mathbb{R}^n$.

   (d) $M = [0, 1] \times [0, 1] \subseteq \mathbb{R}^2$ and $N = [0, 1] \subseteq \mathbb{R}$.

(e) $M = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\}$ and $N = [0, 1] \subseteq \mathbb{R}$.

(f) $M = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\}$ and $N = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + xy + y^2 = 1\}$.

(g) $M = \mathbb{R}^2$ and $N = \mathbb{R}^3$.

4. Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ be a one-to-one continuous function. Show that f must be strictly monotonic in I and f is a homeomorphism between I and $f(I)$.

   If $I \subseteq \mathbb{R}^n$ for $n > 1$ and $f : I \to \mathbb{R}^n$ is continuous and one-to-one, can we still assert that f is homeomorphism between I and $f(I)$?

Problem 4.33 (**True or False**). Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. Let $f : \mathbb{R}^2 \backslash \{(0, 0)\} \to \mathbb{R}$ satisfy $\lim\limits_{x\to 0} f(x, ax^n) = 0$ for all $a \in \mathbb{R}$, $n \in \mathbb{N}$ and $\lim\limits_{y\to 0} f(0, y) = 0$. Then $\lim\limits_{(x,y)\to(0,0)} f(x, y) = 0$.

2. There exists a function $f : \mathbb{R} \to \mathbb{R}$ which is continuous only at three points of $\mathbb{R}$.

3. Let $f : \mathbb{R} \to \mathbb{R}$. Then f is continuous on $\mathbb{R}$ if and only if its graph $\{(x, f(x)) \,|\, x \in \mathbb{R}\}$ is closed in $\mathbb{R}^2$.

Chapter 5

Differentiation of Maps

5.1 Bounded Linear Maps

Definition 5.1. Let X, Y be vector spaces over a common scalar field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A map L from X to Y is said to be ***linear*** if $L(c\boldsymbol{x}_1 + \boldsymbol{x}_2) = cL(\boldsymbol{x}_1) + L(\boldsymbol{x}_2)$ for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in X$ and $c \in \mathbb{F}$. We often write Lx instead of $L(x)$, and the collection of all linear maps from X to Y is denoted by $\mathscr{L}(X, Y)$.

Suppose further that X and Y are normed spaces equipped with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. A linear map $L : X \to Y$ is said to be bounded if

$$\sup_{\|\boldsymbol{x}\|_X = 1} \|L\boldsymbol{x}\|_Y < \infty .$$

The collection of all bounded linear maps from X to Y is denoted by $\mathscr{B}(X, Y)$, and the number $\sup\limits_{\|\boldsymbol{x}\|_X=1} \|L\boldsymbol{x}\|_Y$ is often denoted by $\|L\|_{\mathscr{B}(X,Y)}$.

Example 5.2. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be given by $L\boldsymbol{x} = A\boldsymbol{x}$, where A is an $m \times n$ matrix. Then Example 2.19 shows that $\|L\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}$ is the square root of the largest eigenvalue of $A^{\mathrm{T}}A$ which is certainly a finite number. Therefore, any linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is bounded.

Example 5.3. Recall that $X = \mathscr{C}([a, b]; \mathbb{R})$ and $\|\cdot\|_X = \|\cdot\|_2$ given in Example 2.21 which makes a normed vector space $(X, \|\cdot\|_X)$. Let $\phi \in X$ be given. Define $F : X \to \mathbb{R}$ by

$$F(f) = \int_a^b f(x)\phi(x)\,dx .$$

Then clearly $F \in \mathscr{L}(X, \mathbb{R})$ (the proof is left as an exercise). Moreover, the Cauchy-Schwarz inequality implies that

$$\big|F(f)\big| \leqslant \|f\|_2 \|\phi\|_2 ;$$

thus

$$\sup_{\|f\|_\infty = 1} F(f) \leqslant \|\phi\|_2 < \infty .$$

Therefore, $F \in \mathscr{B}(X, \mathbb{R})$.

Example 5.4. Let $(X, \|\cdot\|_X) = (Y, \|\cdot\|_Y) = (\mathbb{C}, |\cdot|)$, and we consider the bounded linear maps $\mathscr{B}(X, Y)$.

1. Treat $\mathbb{C}$ as a vector space over $\mathbb{C}$. Then $\{1\}$ is a basis of $\mathbb{C}$ and a linear map $L \in \mathscr{L}(\mathbb{C}, \mathbb{C})$ is determined by $L1$ and we have $Lz = zL1$ (thus any linear map from $\mathbb{C}$ to $\mathbb{C}$ is a multiple of a complex number). Moreover,

$$\sup_{|z|=1} |Lz| = |L1| < \infty$$

which shows that $L \in \mathscr{B}(\mathbb{C}, \mathbb{C})$.

2. Treat $\mathbb{C}$ as a vector space over $\mathbb{R}$. Then $\{1, i\}$ is a basis of $\mathbb{C}$ and a linear map $L \in \mathscr{L}(\mathbb{C}, \mathbb{C})$ is determined by $L1$ and Li. In fact, if $L1 = a + bi$ and $Li = c + di$ for some $a, b, c, d \in \mathbb{R}$, then for $x, y \in \mathbb{R}$,

$$L(x + yi) = xL1 + yLi = x(a + bi) + y(c + di) = (ax + cy) + (bx + dy)i .$$

Treating $x + yi$ as a vector $(x, y) \in \mathbb{R}^2$, the map L maps $\begin{bmatrix} x \\ y \end{bmatrix}$ to $\begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. Since

$$\sup_{|x+yi|=1} |L(x + yi)| = \sup_{|x+yi|=1} \big|(ax + cy) + (bx + dy)i\big| = \sup_{\|(x,y)\|_2 = 1} \Big\| \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \Big\|_2 ,$$

the norm of L is the same as the 2-norm of the matrix $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$. Therefore,

$$\|L\|_{\mathscr{B}(\mathbb{C},\mathbb{C})} = \text{ the square root of the largest eigenvalue of } \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} < \infty ;$$

thus $L \in \mathscr{B}(\mathbb{C}, \mathbb{C})$. Since $\mathbb{C}$ over $\mathbb{R}$ is identical to $\mathbb{R}^2$ over $\mathbb{R}$ (from the discussion above), we always treat $\mathbb{C}$ as a vector space over $\mathbb{C}$.

Remark 5.5. By treating $\mathbb{C}$ as a vector space over $\mathbb{C}$ (which, we emphasize again, will always be the case), there is only one linear map from $\mathbb{C}$ to $\mathbb{R}$, the trivial linear map (which sends any vectors to the zero vector). This result is left as an exercise.

Proposition 5.6. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces over a common scalar field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $L \in \mathscr{B}(X, Y)$. Then*

$$\|L\|_{\mathscr{B}(X,Y)} = \sup_{x \neq 0} \frac{\|L\boldsymbol{x}\|_Y}{\|\boldsymbol{x}\|_X} = \inf\big\{M > 0 \,\big|\, \|L\boldsymbol{x}\|_Y \leqslant M\|\boldsymbol{x}\|_X\big\}.$$

In particular, the first equality implies that

$$\|L\boldsymbol{x}\|_Y \leqslant \|L\|_{\mathscr{B}(X,Y)}\|\boldsymbol{x}\|_X \qquad \forall\, \boldsymbol{x} \in X.$$

Proposition 5.7. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces over a common scalar field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $L \in \mathscr{L}(X, Y)$. Then L is continuous on X if and only if $L \in \mathscr{B}(X, Y)$.*

Proof. "$\Rightarrow$" Since L is continuous at $0 \in X$, there exists $\delta > 0$ such that

$$\|L\boldsymbol{x}\|_Y = \|L\boldsymbol{x} - L\mathbf{0}\|_Y < 1 \quad \text{whenever} \quad \|\boldsymbol{x}\|_X < \delta.$$

Then $\big\|L\big(\frac{\delta}{2}\boldsymbol{x}\big)\big\|_Y \leqslant 1$ whenever $\big\|\frac{\delta}{2}\boldsymbol{x}\big\|_X < \delta$; thus by the linearity of L and properties of norms,

$$\|L\boldsymbol{x}\|_Y \leqslant \frac{2}{\delta} \quad \text{whenever} \quad \|\boldsymbol{x}\|_X < 2.$$

Therefore, $\sup\limits_{\|\boldsymbol{x}\|_X=1} \|L\boldsymbol{x}\|_Y \leqslant \frac{2}{\delta}$ which implies that $L \in \mathscr{B}(X, Y)$.

"$\Leftarrow$" If $L \in \mathscr{B}(X, Y)$, then $M = \|L\|_{\mathscr{B}(X,Y)} < \infty$, and

$$\|L\boldsymbol{x}_1 - L\boldsymbol{x}_2\|_Y = \|L(\boldsymbol{x}_1 - \boldsymbol{x}_2)\|_Y \leqslant M\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_X$$

which shows that L is uniformly continuous on X. □

Proposition 5.8. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces over a common scalar field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Then $\big(\mathscr{B}(X, Y), \|\cdot\|_{\mathscr{B}(X,Y)}\big)$ is a normed space. Moreover, if $(Y, \|\cdot\|_Y)$ is a Banach space, so is $\big(\mathscr{B}(X, Y), \|\cdot\|_{\mathscr{B}(X,Y)}\big)$.*

Proof. That $\big(\mathscr{B}(X, Y), \|\cdot\|_{\mathscr{B}(X,Y)}\big)$ is a normed space is left as an exercise. Now suppose that $\big(Y, \|\cdot\|_Y\big)$ is a Banach space. Let $\{L_k\}_{k=1}^{\infty} \subseteq \mathscr{B}(X, Y)$ be a Cauchy sequence. Then by Proposition 5.6, for each $\boldsymbol{x} \in X$ we have

$$\|L_k\boldsymbol{x} - L_\ell\boldsymbol{x}\|_Y = \|(L_k - L_\ell)\boldsymbol{x}\|_Y \leqslant \|L_k - L_\ell\|_{\mathscr{B}(X,Y)}\|\boldsymbol{x}\|_X \to 0 \quad \text{as} \quad k, \ell \to \infty.$$

Therefore, for each $\boldsymbol{x} \in X$ the sequence $\{L_k\boldsymbol{x}\}_{k=1}^{\infty}$ is Cauchy in Y; thus convergent. Suppose that $\lim\limits_{k\to\infty} L_k\boldsymbol{x} = \boldsymbol{y}$. We then establish a map $\boldsymbol{x} \mapsto \boldsymbol{y}$ which we denoted by L; that is, $L\boldsymbol{x} = \boldsymbol{y}$. Then L is linear since if $\boldsymbol{x}_1, \boldsymbol{x}_2 \in X$ and $c \in \mathbb{F}$,

$$L(c\boldsymbol{x}_1 + \boldsymbol{x}_2) = \lim_{k\to\infty} L_k(c\boldsymbol{x}_1 + \boldsymbol{x}_2) = \lim_{k\to\infty} \big(cL_k\boldsymbol{x}_1 + L_k\boldsymbol{x}_2\big) = cL\boldsymbol{x}_1 + L\boldsymbol{x}_2 .$$

Moreover, since $\{L_k\}_{k=1}^{\infty}$ is a Cauchy sequence, by Proposition 2.58 there exists $M > 0$ such that $\|L_k\|_{\mathscr{B}(X,Y)} \leqslant M$ for all $k \in \mathbb{N}$. For each $\boldsymbol{x} \in X$ there exists $N = N_x > 0$ such that

$$\|L_k\boldsymbol{x} - L\boldsymbol{x}\|_Y < 1 \quad \text{ whenever } \quad k \geqslant N_x .$$

Therefore, for $k \geqslant N_x$,

$$\|L\boldsymbol{x}\|_Y < \|L_k\boldsymbol{x}\|_Y + 1 \leqslant \|L_k\|_{\mathscr{B}(X,Y)}\|\boldsymbol{x}\|_X + 1 \leqslant M\|\boldsymbol{x}\|_X + 1$$

which implies that $\sup\limits_{\|\boldsymbol{x}\|_X=1} \|L\boldsymbol{x}\|_Y \leqslant M + 1$; thus $L \in \mathscr{B}(X,Y)$.

Finally, we show that $\lim\limits_{k\to\infty} \|L_k - L\|_{\mathscr{B}(X,Y)} = 0$. Let $\varepsilon > 0$ be given. Since $\{L_k\}_{k=1}^{\infty}$ is a Cauchy sequence, there exists $N > 0$ such that $\|L_k - L_\ell\|_{\mathscr{B}(X,Y)} < \dfrac{\varepsilon}{2}$ whenever $k, \ell \geqslant N$. Then if $k \geqslant N$, for every $\boldsymbol{x} \in X$ we have

$$\|L_k\boldsymbol{x} - L\boldsymbol{x}\|_Y = \lim_{\ell\to\infty} \|L_k\boldsymbol{x} - L_\ell\boldsymbol{x}\|_Y \leqslant \limsup_{\ell\to\infty} \|L_k - L_\ell\|_{\mathscr{B}(X,Y)}\|\boldsymbol{x}\|_X \leqslant \frac{\varepsilon}{2}\|\boldsymbol{x}\|_X ;$$

thus $\|L_k - L\|_{\mathscr{B}(X,Y)} < \varepsilon$ whenever $k \geqslant N$. □

Proposition 5.9. *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ be normed spaces over a common scalar field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $L \in \mathscr{B}(X,Y)$, $K \in \mathscr{B}(Y,Z)$. Then $K \circ L \in \mathscr{B}(X,Z)$, and*

$$\|K \circ L\|_{\mathscr{B}(X,Z)} \leqslant \|K\|_{\mathscr{B}(Y,Z)}\|L\|_{\mathscr{B}(X,Y)} .$$

We often write $K \circ L$ as KL if K and L are linear.

Proof. By the properties of the norm of a bounded linear map,

$$\|K \circ L(x)\|_Z = \|K(Lx)\|_Z \leqslant \|K\|_{\mathscr{B}(Y,Z)}\|Lx\|_Y \leqslant \|K\|_{\mathscr{B}(Y,Z)}\|L\|_{\mathscr{B}(X,Y)}\|x\|_X . \qquad \square$$

From now on, when the domain X and the target Y of a linear map L is clear, we use $\|L\|$ instead of $\|L\|_{\mathscr{B}(X,Y)}$ to simplify the notation.

Theorem 5.10. *Let* $(X, \|\cdot\|_X)$ *and* $(Y, \|\cdot\|_Y)$ *be normed spaces over a common scalar field* $\mathbb{F}$*, where* $\mathbb{F} = \mathbb{R}$ *or* $\mathbb{C}$*, and* X *be finite dimensional. Then every linear map from* X *to* Y *is bounded; that is,* $\mathscr{L}(X, Y) = \mathscr{B}(X, Y)$*.*

Proof. Let $\{\mathbf{e}_k\}_{k=1}^n$ be a basis of X (so that $\dim(X) = n$). Then every $x \in X$ can be expressed as a unique linear combination of $\mathbf{e}_k$'s; that is, for all $\boldsymbol{x} \in X$, there exist unique n numbers $c_1 = c_1(\boldsymbol{x}), \cdots, c_n = c_n(\boldsymbol{x}) \in \mathbb{F}$ such that

$$\boldsymbol{x} = c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n\,.$$

Define an inner product $\langle\cdot,\cdot\rangle$ on X by

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{k=1}^n c_k(\boldsymbol{x})\overline{c_k(\boldsymbol{y})}$$

and let $\|\cdot\|_2$ be the norm induced by this inner product; that is, $\|\boldsymbol{x}\|_2 = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$. That $\langle\cdot,\cdot\rangle$ is indeed an inner product on X is left as an exercise.

Having define $\langle\cdot,\cdot\rangle$, these coefficients c_k's in fact are determined by $c_k(\boldsymbol{x}) = \langle \boldsymbol{x}, \mathbf{e}_k\rangle$, and, by Example 4.29 and the Cauchy-Schwarz inequality, satisfy

$$\big|c_k(\boldsymbol{x})\big| \leqslant \|\boldsymbol{x}\|_2\|\mathbf{e}_k\|_2 \leqslant C\|\boldsymbol{x}\|_X \qquad \forall\, 1 \leqslant k \leqslant n$$

for some constant $C > 0$. As a consequence, if L is a linear map from X to Y, then

$$\begin{aligned}\|L\boldsymbol{x}\|_Y = \big\|L(c_1(\boldsymbol{x})\mathbf{e}_1 + \cdots + c_n(\boldsymbol{x})\mathbf{e}_n)\big\|_Y &\leqslant |c_1(\boldsymbol{x})|\|L\mathbf{e}_1\|_Y + \cdots |c_n(\boldsymbol{x})|\|L\mathbf{e}_n\|_Y \\ &\leqslant C\big(\|L\mathbf{e}_1\|_Y + \|L\mathbf{e}_2\|_Y + \cdots + \|L\mathbf{e}_n\|_Y\big)\|\boldsymbol{x}\|_X \leqslant M\|\boldsymbol{x}\|_X\end{aligned}$$

for some constant $M > 0$; thus $\|L\|_{\mathscr{B}(X,Y)} \leqslant M < \infty$ which shows that $L \in \mathscr{B}(X, Y)$. □

Theorem 5.11. *Let* $\mathrm{GL}(n)$ *be the set of all invertible linear maps on* $\mathbb{R}^n$*; that is,*

$$\mathrm{GL}(n) = \big\{L \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^n)\,\big|\, L \textit{ is one-to-one (and onto)}\big\}.$$

1. *If* $L \in \mathrm{GL}(n)$ *and* $K \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$ *satisfying* $\|K - L\|\|L^{-1}\| < 1$*, then* $K \in \mathrm{GL}(n)$*.*

2. $\mathrm{GL}(n)$ *is an open set of* $\mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$*.*

3. *The mapping* $L \mapsto L^{-1}$ *is continuous on* $\mathrm{GL}(n)$*.*

Proof. 1. Let $\|L^{-1}\| = \dfrac{1}{\alpha}$ and $\|K - L\| = \beta$. Then $\beta < \alpha$; thus for every $\boldsymbol{x} \in \mathbb{R}^n$,

$$\begin{aligned}\alpha\|\boldsymbol{x}\|_{\mathbb{R}^n} &= \alpha\|L^{-1}L\boldsymbol{x}\|_{\mathbb{R}^n} \leqslant \alpha\|L^{-1}\|\|L\boldsymbol{x}\|_{\mathbb{R}^n} = \|L\boldsymbol{x}\|_{\mathbb{R}^n} \leqslant \|(L-K)\boldsymbol{x}\|_{\mathbb{R}^n} + \|K\boldsymbol{x}\|_{\mathbb{R}^n} \\ &\leqslant \beta\|\boldsymbol{x}\|_{\mathbb{R}^n} + \|K\boldsymbol{x}\|_{\mathbb{R}^n}.\end{aligned}$$

As a consequence, $(\alpha - \beta)\|\boldsymbol{x}\|_{\mathbb{R}^n} \leqslant \|K\boldsymbol{x}\|_{\mathbb{R}^n}$ and this implies that $K : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one hence invertible.

2. By 1, we find that if $\|K - L\| < \dfrac{1}{\|L^{-1}\|}$, then $K \in \mathrm{GL}(n)$. Then $B\Big(L, \dfrac{1}{\|L^{-1}\|}\Big) \subseteq \mathrm{GL}(n)$ if $L \in \mathrm{GL}(n)$. Therefore, $\mathrm{GL}(n)$ is open.

3. Let $L \in \mathrm{GL}(n)$ and $\varepsilon > 0$ be given. Choose $\delta = \min\Big\{\dfrac{1}{2\|L^{-1}\|}, \dfrac{\varepsilon}{2\|L^{-1}\|^2}\Big\}$. Then $\delta > 0$, and $K \in \mathrm{GL}(n)$ whenever $\|K - L\| < \delta$. Since $K^{-1} - L^{-1} = K^{-1}(L - K)L^{-1}$, we find that if $\|K - L\| < \delta$,
$$\|K^{-1}\| - \|L^{-1}\| \leqslant \|K^{-1} - L^{-1}\| \leqslant \|K^{-1}\|\|K - L\|\|L^{-1}\| < \frac{1}{2}\|K^{-1}\|$$
which implies that $\|K^{-1}\| < 2\|L^{-1}\|$. Therefore, if $\|K - L\| < \delta$,
$$\|K^{-1} - L^{-1}\| \leqslant \|K^{-1}\|\|K - L\|\|L^{-1}\| < 2\|L^{-1}\|^2\delta \leqslant \varepsilon. \qquad \square$$

5.2 Definition of Derivatives

Definition 5.12. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed spaces over a common scalar field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A map $f : A \subseteq X \to Y$ is said to be ***differentiable*** at $a \in A$ if there exists a map in $\mathscr{B}(X, Y)$, denoted by $(Df)(a)$ and called the ***derivative*** of f at a, such that
$$\lim_{x \to a} \frac{\big\|f(x) - f(a) - (Df)(a)(x - a)\big\|_Y}{\|x - a\|_X} = 0,$$
where $(Df)(a)(x - a)$ denotes the value of the bounded linear map $(Df)(a)$ applied to the vector $x - a \in X$ (so $(Df)(a)(x - a) \in Y$). In other words, f is differentiable at $a \in A$ if there exists $L \in \mathscr{B}(X, Y)$ such that
$$\forall\, \varepsilon > 0, \exists\, \delta > 0 \ni \|f(x) - f(a) - L(x - a)\|_Y \leqslant \varepsilon\|x - a\|_X \text{ whenever } x \in B(a, \delta) \cap A.$$
If f is differentiable at each point of A, we say that f is differentiable on A.

Remark 5.13. If f is differentiable on A, then for each $x \in A$, $(Df)(x)$ is a bounded linear map from X to Y, but Df in general is not linear in x.

Remark 5.14. The condition

$$\lim_{x \to a} \frac{\big\|f(x) - f(a) - (Df)(a)(x-a)\big\|_Y}{\|x-a\|_X} = 0\,,$$

is sometimes written as

$$f(x) = f(a) + (Df)(a)(x-a) + \mathscr{o}(\|x-a\|_X) \quad \text{as} \quad x \to a\,,$$

where $f = g + \mathscr{o}(h)$ as $x \to a$ is a short-hand notation for $\lim\limits_{x \to a} \dfrac{\|f-g\|}{h} = 0$.

Remark 5.15. Let $a \in A$ and v be a unit vector in $(X, \|\cdot\|_X)$ such that $a + tv \in A$ for all $t \in [0,1]$. If $f : A \to \mathbb{R}$ is differentiable at a, then

$$\lim_{t \to 0^+} \frac{\big|f(a+tv) - f(a) - (Df)(a)(tv)\big|}{\|(a+tv) - a\|_X} = 0\,.$$

Since $(Df)(a)(tv) = t(Df)(a)(v)$ and $\|tv\|_X = t$ (since $t > 0$), the identity above implies that

$$\lim_{t \to 0^+} \Big|\frac{f(a+tv) - f(a)}{t} - (Df)(a)(v)\Big| = 0$$

or equivalently,

$$\lim_{t \to 0^+} \frac{f(a+tv) - f(a)}{t} = (Df)(a)(v)\,.$$

In Calculus, the limit on the left-hand side is the ***directional derivative of*** f ***at*** a ***in direction*** v and is usually denoted by $(D_v f)(a)$; thus the quantity $(Df)(a)(v)$ is a generalization of the directional derivative.

Example 5.16. Let $f : (a,b) \to \mathbb{R}$ be differentiable at $c \in (a,b)$. Then there exists $L \in \mathscr{B}(\mathbb{R}, \mathbb{R})$ such that

$$\lim_{x \to c} \frac{\big|f(x) - f(c) - L(x-c)\big|}{|x-c|} = 0\,.$$

Since $L \in \mathscr{B}(\mathbb{R}, \mathbb{R})$, there exists a real number m such that $L(x) = mx$ for all $x \in \mathbb{R}$; thus the identity above implies that

$$\lim_{x \to c} \frac{f(x) - f(c) - m(x-c)}{x-c} = 0$$

or equivalently,

$$\lim_{x \to c} \frac{f(x) - f(c)}{x-c} = m\,.$$

In other words, a function $f : (a,b) \to \mathbb{R}$ is differentiable at $c \in (a,b)$ if and only if the limit $\lim\limits_{x \to c} \dfrac{f(x) - f(c)}{x-c}$ exists. The limit is usually denoted by $f'(c)$, and we identify the linear map L with the real number $f'(c)$ using the relation $L(x) = f'(c)x$.

Example 5.17. Let $f : (0,\infty) \to \mathbb{R}$ be given by $f(x) = \dfrac{1}{x}$. Then f is differentiable at any $a \in (0,\infty)$ since $(Df)(a) : \mathbb{R} \to \mathbb{R}$ is the linear map given by

$$(Df)(a)(x) = -\frac{1}{a^2} \cdot x\,.$$

To see this, we observe that

$$\begin{aligned}\lim_{x\to a} \frac{\left|\frac{1}{x} - \frac{1}{a} - \frac{-1}{a^2}(x-a)\right|}{|x-a|} &= \lim_{x\to a} \frac{\left|\frac{a^2 - xa + x^2 - xa}{xa^2}\right|}{|x-a|} = \lim_{x\to a} \frac{a^2 - 2xa + x^2}{xa^2|x-a|} \\ &= \lim_{x\to a} \frac{|x-a|}{xa^2} = 0\,.\end{aligned}$$

Example 5.18. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed spaces. Then every bounded linear map $L : X \to Y$ is differentiable. In fact, $(DL)(a) = L$ for all $a \in X$ since

$$\lim_{x\to a} \frac{\|Lx - La - L(x-a)\|_Y}{\|x-a\|_X} = 0\,.$$

Example 5.19. Recall that $\mathcal{M}_{n\times n}$ denotes the collection of all $n \times n$ real matrices. Equip it with 2-norm and let $f : \mathcal{M}_{n\times n} \to \mathcal{M}_{n\times n}$ be given by $f(L) = L^2$. Then for $K, L \in \mathcal{M}_{n\times n}$,

$$f(K) - f(L) = K^2 - L^2 = L(K-L) + (K-L)L + (K-L)^2\,.$$

This motivates us to define $(Df)(L)$ by $(Df)(L)(H) = LH + HL$ so that

$$\|f(K) - f(L) - (Df)(L)(K-L)\|_2 \leqslant \|K-L\|_2^2\,;$$

which shows

$$\lim_{K\to L} \frac{\|f(K) - f(L) - (Df)(L)(K-L)\|_2}{\|K-L\|_2} = 0\,.$$

Therefore, f is differentiable at every $L \in \mathcal{M}_{n\times n}$.

Example 5.20. Let $f : \mathrm{GL}(n) \to \mathrm{GL}(n)$ be given by $f(L) = L^{-1}$, where $\mathrm{GL}(n)$ is defined in Theorem 5.11. Then f is differentiable at any "point" $L \in \mathrm{GL}(n)$ with derivative $(Df)(L) \in \mathscr{B}(\mathrm{GL}(n), \mathrm{GL}(n))$ given by

$$(Df)(L)(K) = -L^{-1}KL^{-1} \qquad \text{for all } K \in \mathrm{GL}(n)\,.$$

To see this, for $K, L \in \mathrm{GL}(n)$,

$$\begin{aligned}\big\|f(K) - f(L) + L^{-1}(K-L)L^{-1}\big\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))} &= \big\|(L^{-1} - K^{-1})(K-L)L^{-1}\big\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))} \\ \leqslant \|L^{-1} - K^{-1}\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))}&\|K-L\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))}\|L^{-1}\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))}\,;\end{aligned}$$

thus if $K \neq L$,

$$\frac{\|f(K)-f(L)+L^{-1}(K-L)L^{-1}\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))}}{\|K-L\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))}} \leqslant \|L^{-1}\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))}\|L^{-1}-K^{-1}\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))}.$$

By the fact that the map $L \mapsto L^{-1}$ is continuous on $\mathrm{GL}(n)$ (Theorem 5.11), we find that

$$\lim_{K\to L}\frac{\|f(K)-f(L)+L^{-1}(K-L)L^{-1}\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))}}{\|K-L\|_{\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))}} = 0.$$

Example 5.21. Recall the setting in Example 4.11 that $X = \mathscr{C}([a,b];\mathbb{R})$, $\|\cdot\|_X = \|\cdot\|_2$, and $\mathrm{I} : X \to \mathbb{R}$ given by

$$\mathrm{I}(f) = \int_a^b \big|f(x)\big|^2\,dx.$$

Then I is differentiable at every $f \in X$ since if $(D\mathrm{I})(f)(h) \equiv 2\int_a^b f(x)h(x)\,dx$, Example 5.3 shows that $(D\mathrm{I})(f) \in \mathscr{B}(X,\mathbb{R})$ and

$$\begin{aligned}\big|\mathrm{I}(g)-\mathrm{I}(f)-(D\mathrm{I})(f)(g-f)\big| &= \Big|\int_a^b \Big[|g(x)|^2-|f(x)|^2-2f(x)\big(g(x)-f(x)\big)\Big]\,dx\Big| \\ &= \int_a^b \big[g(x)-f(x)\big]^2\,dx = \|f-g\|_2^2\end{aligned}$$

which implies that

$$\lim_{g\to f}\frac{\big|\mathrm{I}(g)-\mathrm{I}(f)-(D\mathrm{I})(f)(g-f)\big|}{\|g-f\|_2} = 0.$$

Example 5.22. The function $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x,y) = x^2+y^2$ is differentiable at every $(a,b) \in \mathbb{R}^2$. In fact, $(Df)(a,b)(h,k)$, the linear map $(Df)(a,b)$ acting on the vector (h,k) is $(Df)(a,b)(h,k) = 2ah+2bk$, since

$$\begin{aligned}&\lim_{(x,y)\to(a,b)}\frac{\big|f(x,y)-f(a,b)-(Df)(a,b)(x-a,y-b)\big|}{\|(x,y)-(a,b)\|} \\ &\quad= \lim_{(h,k)\to(0,0)}\frac{\big|f(a+h,b+k)-f(a,b)-(Df)(a,b)(h,k)\big|}{\|(h,k)\|} \\ &\quad= \lim_{(h,k)\to(0,0)}\frac{\big|(a+h)^2+(b+k)^2-a^2-b^2-2ah-2bk\big|}{\sqrt{h^2+k^2}} \\ &\quad= \lim_{(h,k)\to(0,0)}\frac{h^2+k^2}{\sqrt{h^2+k^2}} = \lim_{(h,k)\to(0,0)}\sqrt{h^2+k^2} = 0.\end{aligned}$$

On the other hand, Example 5.5 shows that the function $g : \mathbb{C} \to \mathbb{R}$ given by $g(z) = |z|^2$ is differentiable only if $g'(z) \equiv (Dg)(z) = 0$. Therefore, if g is differentiable at z_0, then

$$0 = \lim_{z \to z_0} \frac{\big||z|^2 - |z_0|^2\big|}{|z - z_0|} = \lim_{h \to 0} \frac{\big|(z_0 + h) \cdot \overline{z_0 + h} - z_0\overline{z_0}\big|}{|h|} = \lim_{h \to 0} \Big|\bar{z}_0 - z_0 \frac{\bar{h}}{h}\Big|;$$

thus $z_0 = 0$ since $\lim\limits_{h \to 0} \dfrac{\bar{h}}{h}$ does not exist.

By treating $\mathbb{R}$ as a subset of $\mathbb{C}$, we treat g as a function from $\mathbb{C}$ to $\mathbb{C}$ (note that there are more maps in $\mathscr{B}(\mathbb{C}, \mathbb{C})$ so in principle it is easier to have differentiable functions from $\mathbb{C}$ to $\mathbb{C}$). Assume the contrary that $g'(z_0) \equiv (Dg)(z_0) \in \mathscr{B}(\mathbb{C}, \mathbb{C})$ exists, then

$$\begin{aligned} 0 &= \lim_{z \to z_0} \frac{\big|g(z) - g(z_0) - g'(z_0)(z - z_0)\big|}{|z - z_0|} = \lim_{h \to 0} \frac{\big|g(z_0 + h) - g(z_0) - g'(z_0)h\big|}{|h|} \\ &= \lim_{h \to 0} \frac{\big|(z_0 + h)\overline{(z_0 + h)} - z_0\overline{z_0} - g'(z_0)h\big|}{|h|} = \lim_{h \to 0} \frac{\big|h\bar{z}_0 + z_0\bar{h} + |h|^2 - g'(z_0)h\big|}{|h|} \\ &= \lim_{h \to 0} \Big|\bar{z}_0 - g'(z_0) + z_0 \frac{\bar{h}}{|h|}\Big|; \end{aligned}$$

thus $\lim\limits_{h \to 0} \Big(\bar{z}_0 - g'(z_0) + z_0 \dfrac{\bar{h}}{|h|}\Big) = 0$ or equivalently, $z_0 \Big(\lim\limits_{h \to 0} \dfrac{\bar{h}}{|h|}\Big) = g'(z_0) - \bar{z}_0$. If $z_0 \neq 0$, then the fact that $\lim\limits_{h \to 0} \dfrac{\bar{h}}{|h|}$ does not exist shows that the identity above cannot be true; thus g is not differentiable at $z_0 \neq 0$. On the other hand, if $z_0 = 0$, the choice of $g'(z_0) = 0$ makes the identity valid. Therefore, g is differentiable only at 0 and $g'(0) = 0$. This agrees with the observation by treating g as a function from $\mathbb{C}$ to $\mathbb{R}$.

Note that by writing $z = x + iy$, we indeed have $g(x + iy) = x^2 + y^2$; thus $f(x, y) = g(x + iy)$. Even though f is differentiable at every point in $\mathbb{R}^2$, g is not. The reason behind this is that there are "much more" bounded linear maps in $\mathscr{B}(\mathbb{R}^2, \mathbb{R})$ than bounded linear maps in $\mathscr{B}(\mathbb{C}, \mathbb{R})$ or $\mathscr{B}(\mathbb{C}, \mathbb{C})$ so that it is easier to make a function from $\mathbb{R}^2 \to \mathbb{R}$ differentiable.

Theorem 5.23. *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed vector spaces, A be a subset of X, and $f : A \to Y$ be differentiable at a. If $a \in \mathring{A}$, then $(Df)(a)$ is uniquely determined by f.*

Proof. Suppose $L_1, L_2 \in \mathscr{B}(X, Y)$ are derivatives of f at a. Let $\varepsilon > 0$ be given and $\mathrm{e} \in X$ be a unit vector; that is, $\|\mathrm{e}\|_X = 1$. Since a is an interior point of A, there exists $r > 0$ such that $B(a, r) \subseteq A$. By Definition 5.12, there exists $0 < \delta < r$ such that

$$\frac{\|f(x) - f(a) - L_1(x - a)\|_Y}{\|x - a\|_X} < \frac{\varepsilon}{2} \qquad \text{and} \qquad \frac{\|f(x) - f(a) - L_2(x - a)\|_Y}{\|x - a\|_X} < \frac{\varepsilon}{2}$$

if $0 < \|x-a\|_X < \delta$. Letting $x = a + \lambda\mathrm{e}$ with $0 < |\lambda| < \delta$, we have

$$\begin{aligned}\|L_1\mathrm{e} - L_2\mathrm{e}\|_Y &= \frac{1}{|\lambda|}\|L_1(x-a) - L_2(x-a)\|_Y \\ &\leqslant \frac{1}{|\lambda|}\big(\big\|f(x)-f(a)-L_1(x-a)\big\|_Y + \big\|f(x)-f(a)-L_2(x-x_2)\big\|_Y\big) \\ &= \frac{\big\|f(x)-f(a)-L_1(x-a)\big\|_Y}{\|x-a\|_X} + \frac{\big\|f(x)-f(a)-L_2(x-a)\big\|_Y}{\|x-a\|_X} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we conclude that $L_1\mathrm{e} = L_2\mathrm{e}$ for all unit vectors $\mathrm{e} \in X$ which guarantees that $L_1 = L_2$ (since if $x \neq 0$, $L_1 x = \|x\|_X L_1\Big(\frac{x}{\|x\|_X}\Big) = \|x\|_X L_2\Big(\frac{x}{\|x\|_X}\Big) = L_2 x$). ▫

Example 5.24. $(Df)(a)$ may not be unique if the domain of f is not open. For example, let $A = \big\{(x,y)\,\big|\,0 \leqslant x \leqslant 1, y = 0\big\}$ be a subset of $\mathbb{R}^2$, and $f : A \to \mathbb{R}$ be given by $f(x,y) = 0$. Fix $a = (h,0) \in A$, then both of the linear map

$$L_1(x,y) = 0 \quad \text{and} \quad L_2(x,y) = hy \qquad \forall\,(x,y) \in \mathbb{R}^2$$

satisfy Definition 5.12 since

$$\lim_{(x,0)\to(h,0)} \frac{\big|f(x,0)-f(h,0)-L_1(x-h,0)\big|}{\big\|(x,0)-(h,0)\big\|_{\mathbb{R}^2}} = \lim_{(x,0)\to(h,0)} \frac{\big|f(x,0)-f(h,0)-L_2(x-h,0)\big|}{\big\|(x,0)-(h,0)\big\|_{\mathbb{R}^2}} = 0\,.$$

Remark 5.25. Let $U \subseteq \mathbb{R}^n$ be an open set and suppose that $f : U \to \mathbb{R}^m$ is differentiable on U. Then $Df : U \to \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$. Treating Df as a map from U to the normed space $\big(\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m), \|\cdot\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}\big)$, and suppose that Df is also differentiable on U. Then the derivative of Df, denoted by $D^2 f$, is a map from U to $\mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n,\mathbb{R}^m))$. In other words, for each $a \in U$, $(D^2 f)(a) \in \mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n,\mathbb{R}^m))$ satisfying

$$\lim_{x\to a} \frac{\big\|(Df)(x) - (Df)(a) - (D^2 f)(a)(x-a)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}}{\|x-a\|_{\mathbb{R}^n}} = 0\,,$$

here $(D^2 f)(a)$ is bounded linear map from $\mathbb{R}^n$ to $\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)$; thus $(D^2 f)(a)(x-a) \in \mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)$.

Definition 5.26. Let $\{\mathrm{e}_k\}_{k=1}^n$ be the standard basis of $\mathbb{R}^n$, $U \subseteq \mathbb{R}^n$ be an open set, $a \in U$ and $f : U \to \mathbb{R}$ be a function. The partial derivative of f at a in the direction e_j, denoted by $\dfrac{\partial f}{\partial x_j}(a)$, is the limit

$$\lim_{h\to 0}\frac{f(a+he_j)-f(a)}{h}$$

if it exists. In other words, if $a=(a_1,\cdots,a_n)$, then

$$\frac{\partial f}{\partial x_j}(a)=\lim_{h\to 0}\frac{f(a_1,\cdots,a_{j-1},a_j+h,a_{j+1},\cdots,a_n)-f(a_1,\cdots,a_n)}{h}.$$

Theorem 5.27. *Suppose $U\subseteq\mathbb{R}^n$ is an open set and $f:U\to\mathbb{R}^m$ is differentiable at $a\in U$. Then the partial derivatives $\dfrac{\partial f_i}{\partial x_j}(a)$ exists for all $i=1,\cdots m$ and $j=1,\cdots n$, and the matrix representation of the linear map $Df(a)$ with respect to the standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$ is given by*

$$\big[(Df)(a)\big]=\begin{bmatrix}\dfrac{\partial f_1}{\partial x_1}(a) & \cdots & \dfrac{\partial f_1}{\partial x_n}(a)\\ \vdots & \ddots & \vdots\\ \dfrac{\partial f_m}{\partial x_1}(a) & \cdots & \dfrac{\partial f_m}{\partial x_n}(a)\end{bmatrix}\qquad \text{or}\qquad \big[(Df)(a)\big]_{ij}=\frac{\partial f_i}{\partial x_j}(a).$$

Proof. Since U is open and $a\in U$, there exists $r>0$ such that $B(a,r)\subseteq U$. By the differentiability of f at a, there is $L\in\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)$ such that for any given $\varepsilon>0$, there exists $0<\delta<r$ such that

$$\|f(x)-f(a)-L(x-a)\|_{\mathbb{R}^m}\leqslant\varepsilon\|x-a\|_{\mathbb{R}^n}\qquad\text{whenever}\qquad x\in B(a,\delta).$$

In particular, for each $i=1,\cdots,m$,

$$\Big|\frac{f_i(a+he_j)-f_i(a)}{h}-(Le_j)_i\Big|\leqslant\Big\|\frac{f(a+he_j)-f(a)}{h}-Le_j\Big\|_{\mathbb{R}^m}\leqslant\varepsilon\quad\forall\,0<|h|<\delta,h\in\mathbb{R},$$

where $(Le_j)_i$ denotes the i-th component of Le_j in the standard basis. As a consequence, for each $i=1,\cdots,m$,

$$\lim_{h\to 0}\frac{f_i(a+he_j)-f_i(a)}{h}=(Le_j)_i\text{ exists}$$

and by definition, we must have $(Le_j)_i=\dfrac{\partial f_i}{\partial x_j}(a)$. Therefore, $L_{ij}=\dfrac{\partial f_i}{\partial x_j}(a)$. □

Definition 5.28. Let $U\subseteq\mathbb{R}^n$ be an open set, and $f:U\to\mathbb{R}^m$. The matrix

$$(Jf)(x)\equiv\begin{bmatrix}\dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n}\\ \vdots & \ddots & \vdots\\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n}\end{bmatrix}(x)\equiv\begin{bmatrix}\dfrac{\partial f_1}{\partial x_1}(x) & \cdots & \dfrac{\partial f_1}{\partial x_n}(x)\\ \vdots & \ddots & \vdots\\ \dfrac{\partial f_m}{\partial x_1}(x) & \cdots & \dfrac{\partial f_m}{\partial x_n}(x)\end{bmatrix}$$

is called the ***Jacobian matrix*** of f at x (if each entry exists). If $n = m$, the determinant of $(Jf)(x)$ is called the ***Jacobian*** of f at x.

Remark 5.29. A function f might not be differential even if the Jacobian matrix Jf exists; however, if f is differentiable at x_0, then $(Df)(x)$ can be represented by $(Jf)(x)$; that is, $[(Df)(x)] = (Jf)(x)$.

Example 5.30. Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be given by $f(x_1, x_2) = (x_1^2, x_1^3 x_2, x_1^4 x_2^2)$. Suppose that f is differentiable at $x = (x_1, x_2)$, then

$$\big[(Df)(x)\big] = \begin{bmatrix} 2x_1 & 0 \\ 3x_1^2 x_2 & x_1^3 \\ 4x_1^3 x_2^2 & 2x_1^4 x_2 \end{bmatrix}.$$

Example 5.31. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

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

Then $\dfrac{\partial f}{\partial x}(0,0) = \dfrac{\partial f}{\partial y}(0,0) = 0$; thus if f is differentiable at $(0,0)$, then $[(Df)(0,0)] = \begin{bmatrix} 0 & 0 \end{bmatrix}$. However,

$$\Big| f(x,y) - f(0,0) - \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \Big| = \frac{|xy|}{x^2+y^2} = \frac{|xy|}{(x^2+y^2)^{\frac{3}{2}}} \sqrt{x^2+y^2};$$

thus f is not differentiable at $(0,0)$ since $\dfrac{|xy|}{(x^2+y^2)^{\frac{3}{2}}}$ cannot be arbitrarily small even if x^2+y^2 is small.

Example 5.32. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = \begin{cases} x & \text{if } y = 0, \\ y & \text{if } x = 0, \\ 1 & \text{otherwise}. \end{cases}$$

Then $\dfrac{\partial f}{\partial x}(0,0) = \lim\limits_{h\to 0} \dfrac{f(h,0) - f(0,0)}{h} = \lim\limits_{h \to 0} \dfrac{h}{h} = 1$. Similarly, $\dfrac{\partial f}{\partial y}(0,0) = 1$; thus if f is differentiable at $(0,0)$, then $[(Df)(0,0)] = \begin{bmatrix} 1 & 1 \end{bmatrix}$. However,

$$\Big| f(x,y) - f(0,0) - \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \Big| = \big| f(x,y) - (x+y) \big|;$$

thus if $xy \neq 0$,

$$\big|f(x,y)-(x+y)\big| = |1-x-y| \not\to 0 \quad \text{as} \quad (x,y)\to(0,0), xy\neq 0\,.$$

Therefore, f is not differentiable at $(0,0)$.

Definition 5.33. Let $U \subseteq \mathbb{R}^n$ be an open set. The derivative of a scalar function $f : U \to \mathbb{R}$ is called the ***gradient*** of f and is denoted by grad$f$ or ∇f.

5.3 Continuity of Differentiable Maps

Theorem 5.34. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, $U \subseteq X$ be open, and $f : U \to Y$ be differentiable at $a \in U$. Then f is continuous at a.*

Proof. Since f is differentiable at a, there exists $L \in \mathscr{B}(X,Y)$ such that

$$\exists\, \delta > 0 \ni \big\|f(x)-f(a)-L(x-a)\big\|_Y \leqslant \|x-a\|_X \quad \forall\, x \in B(a,\delta)\,.$$

As a consequence,

$$\big\|f(x)-f(a)\big\|_Y \leqslant \big(\|L\|+1\big)\|x-a\|_X \quad \forall\, x \in B(a,\delta)\,; \tag{5.3.1}$$

thus $\lim\limits_{x\to a}\|f(x)-f(a)\|_Y = 0$. □

Remark 5.35. In fact, if f is differentiable at x_0, then f satisfies the "local Lipschitz property"; that is,

$$\exists\, M = M(x_0) > 0 \text{ and } \delta = \delta(x_0) > 0 \ni \|f(x)-f(x_0)\|_Y \leqslant M\|x-x_0\|_X \;\text{ if }\; \|x-x_0\|_X < \delta$$

since we can choose $M = \|L\|+1$ and $\delta = \delta_1$ (see (5.3.1)).

Example 5.36. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given in Example 5.31. We have shown that f is not differentiable at $(0,0)$. In fact, f is not even continuous at $(0,0)$ since when approaching the origin along the straight line $x_2 = mx_1$,

$$\lim_{(x_1,mx_1)\to(0,0)} f(x_1,mx_1) = \lim_{x_1\to 0}\frac{mx_1^2}{(m^2+1)x_1^2} = \frac{m^2}{m^2+1} \neq f(0,0) \text{ if } m\neq 0\,.$$

Example 5.37. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given in Example 5.32. Then f is not continuous at $(0,0)$; thus not differentiable at $(0,0)$.

Example 5.38. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Then $f_x(0,0) = 1$ and $f_y(0,0) = 0$. However,

$$\frac{\left| f(x,y) - f(0,0) - \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right|}{\sqrt{x^2+y^2}} = \frac{|x|y^2}{(x^2+y^2)^{\frac{3}{2}}} \not\to 0 \quad \text{as} \quad (x,y) \to (0,0).$$

Therefore, f is not differentiable at $(0,0)$. On the other hand, f is continuous at $(0,0)$ since

$$\big|f(x,y) - f(0,0)\big| = \big|f(x,y)\big| \leqslant |x| \to 0 \quad \text{as} \quad (x,y) \to (0,0).$$

5.4 Conditions for Differentiability

Proposition 5.39. *Let $U \subseteq \mathbb{R}^n$ be open, $a \in U$, and $f = (f_1, \cdots, f_m) : U \to \mathbb{R}^m$. Then f is differentiable at a if and only if f_i is differentiable at a for all $i = 1, \cdots, m$. In other words, for vector-valued functions defined on an open subset of $\mathbb{R}^n$,*

$$\textit{Componentwise differentiable} \Leftrightarrow \textit{Differentiable.}$$

Proof. By the definition of differentiability and Proposition 2.49,

$$\begin{aligned}
& f \text{ is differentiable at } a \\
& \Leftrightarrow (\exists\, L \in \mathcal{M}_{m\times n})\Big(\lim_{x\to a} \frac{\|f(x) - f(a) - L(x-a)\|_{\mathbb{R}^m}}{\|x-a\|_{\mathbb{R}^n}} = 0\Big) \\
& \Leftrightarrow (\exists\, L \in \mathcal{M}_{m\times n})\Big(\lim_{x\to a} \Big\| \frac{f(x)-f(a)}{\|x-a\|_{\mathbb{R}^n}} - L\Big(\frac{[x-a]}{\|x-a\|_{\mathbb{R}^n}}\Big)\Big\|_{\mathbb{R}^m} = 0\Big) \\
& \Leftrightarrow (\exists\, L \in \mathcal{M}_{m\times n})(\forall\, 1 \leqslant i \leqslant n)\Big(\lim_{x\to a} \Big| \frac{f_i(x)-f_i(a)}{\|x-a\|_{\mathbb{R}^n}} - (\mathrm{e}_i^{\mathrm{T}} L)\Big(\frac{[x-a]}{\|x-a\|_{\mathbb{R}^n}}\Big)\Big| = 0\Big) \\
& \Leftrightarrow (\exists\, L \in \mathcal{M}_{m\times n})(\forall\, 1 \leqslant i \leqslant n)\Big(\lim_{x\to a} \frac{\big|f_i(x)-f_i(a) - (\mathrm{e}_i^{\mathrm{T}} L)(x-a)\big|}{\|x-a\|_{\mathbb{R}^n}} = 0\Big) \\
& \Leftrightarrow (\exists\, L_i \in \mathcal{M}_{m\times 1})\Big(\lim_{x\to a} \frac{\big|f_i(x)-f_i(a) - L_i(x-a)\big|}{\|x-a\|_{\mathbb{R}^n}} = 0\Big) \\
& \Leftrightarrow f_i \text{ is differentiable at } a \text{ for each } 1 \leqslant i \leqslant n.
\end{aligned}$$

□

Theorem 5.40. *Let $U \subseteq \mathbb{R}^n$ be open, $a \in U$, and $f : U \to \mathbb{R}$. If each entry of the Jacobian matrix $\left[\dfrac{\partial f}{\partial x_1} \cdots \dfrac{\partial f}{\partial x_n}\right]$ of f*

1. *exists in a neighborhood of a, and*

2. *is continuous at a except perhaps one entry.*

Then f is differentiable at a.

Proof. W.L.O.G. we can assume that $\dfrac{\partial f}{\partial x_i}$ is continuous at a for $1 \leqslant i < n$. Let $\{\mathrm{e}_j\}_{j=1}^n$ be the standard basis of $\mathbb{R}^n$, and $\varepsilon > 0$ be given. Since $\dfrac{\partial f}{\partial x_i}$ is continuous at a for $i = 1, \cdots, n-1$,

$$\exists\, \delta_i > 0 \ni \Big|\frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(a)\Big| < \frac{\varepsilon}{\sqrt{n}} \quad \text{whenever} \quad \|x-a\|_{\mathbb{R}^n} < \delta_i\,.$$

On the other hand, by the definition of the partial derivatives,

$$\exists\, \delta_n > 0 \ni \Big|\frac{f(a+h\mathrm{e}_n)-f(a)}{h} - \frac{\partial f}{\partial x_n}(a)\Big| \leqslant \frac{\varepsilon}{\sqrt{n}} \quad \text{whenever} \quad |h| < \delta_n\,.$$

Let $k = x - a$ and $\delta = \min\big\{\delta_1, \cdots, \delta_n\big\}$. Then

$$\begin{aligned}
&\Big|f(x) - f(a) - \Big[\frac{\partial f}{\partial x_1}(a)(x_1-a_1) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n-a_n)\Big]\Big| \\
&\quad = \Big|f(a+k) - f(a) - \frac{\partial f}{\partial x_1}(a)k_1 - \cdots - \frac{\partial f}{\partial x_n}(a)k_n\Big| \\
&\quad = \Big|f(a_1+k_1, \cdots, a_n+k_n) - f(a_1, \cdots, a_n) - \frac{\partial f}{\partial x_1}(a)k_1 - \cdots - \frac{\partial f}{\partial x_n}(a)k_n\Big| \\
&\quad \leqslant \Big|f(a_1+k_1, \cdots, a_n+k_n) - f(a_1, a_2+k_2, \cdots, a_n+k_n) - \frac{\partial f}{\partial x_1}(a)k_1\Big| \\
&\qquad + \Big|f(a_1, a_2+k_2, \cdots, a_n+k_n) - f(a_1, a_2, a_3+k_3, \cdots, a_n+k_n) - \frac{\partial f}{\partial x_2}(a)k_2\Big| \\
&\qquad + \cdots + \Big|f(a_1, \cdots, a_{n-1}, a_n+k_n) - f(a_1, \cdots, a_n) - \frac{\partial f}{\partial x_n}(a)k_n\Big|.
\end{aligned}$$

By the mean value theorem,

$$\begin{aligned}
&f(a_1, \cdots, a_{j-1}, a_j+k_j, \cdots, a_n+k_n) - f(a_1, \cdots, a_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n) \\
&\quad = k_j \frac{\partial f}{\partial x_j}(a_1, \cdots, a_{j-1}, a_j+\theta_j k_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n)
\end{aligned}$$

for some $0 < \theta_j < 1$; thus for $j = 1, \cdots, n-1$, if $\|x-a\|_{\mathbb{R}^n} = \|k\|_{\mathbb{R}^n} < \delta$,

$$\begin{aligned}
&\Big|f(a_1, \cdots, a_{j-1}, a_j+k_j, \cdots, a_n+k_n) - f(a_1, \cdots, a_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n) - \frac{\partial f}{\partial x_j}(a)k_j\Big| \\
&\quad = \Big|\frac{\partial f}{\partial x_j}(a_1, \cdots, a_{j-1}, a_j+\theta_j k_j, a_{j+1}+k_{j+1}, \cdots, a_n+k_n) - \frac{\partial f}{\partial x_j}(a)\Big||k_j| \leqslant \frac{\varepsilon}{\sqrt{n}}|k_j|\,.
\end{aligned}$$

Moreover, if $\|x-a\|_{\mathbb{R}^n} < \delta$, then $|k_n| \leqslant \|k\|_{\mathbb{R}^n} = \|x-a\|_{\mathbb{R}^n} < \delta \leqslant \delta_n$; thus

$$\Big|f(a_1,\cdots,a_{n-1},a_n+k_n) - f(a_1,\cdots,a_n) - \frac{\partial f}{\partial x_n}(a)k_n\Big| \leqslant \frac{\varepsilon}{\sqrt{n}}|k_n|.$$

As a consequence, if $\|x-a\|_{\mathbb{R}^n} < \delta$, by Cauchy's inequality,

$$\begin{aligned}\Big|f(x) - f(a) - \Big[\frac{\partial f}{\partial x_1}(a)(x_1-a_1) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n-a_n)\Big]\Big| \\ \leqslant \frac{\varepsilon}{\sqrt{n}}\sum_{j=1}^{n}|k_j| \leqslant \varepsilon\|k\|_{\mathbb{R}^n} = \varepsilon\|x-a\|_{\mathbb{R}^n}\end{aligned}$$

which implies that f is differentiable at a. □

Remark 5.41. When two or more components of the Jacobian matrix $\Big[\dfrac{\partial f}{\partial x_1} \ \cdots \ \dfrac{\partial f}{\partial x_n}\Big]$ of a scalar function f are discontinuous at a point $a \in U$, in general f is not differentiable at a. For example, both components of the Jacobian matrix of the functions given in Example 5.31, 5.32, 5.38 are discontinuous at $(0,0)$, and these functions are not differentiable at $(0,0)$.

Example 5.42. Let $U = \mathbb{R}^2 \backslash \{(x,0) \in \mathbb{R}^2 \,|\, x \geqslant 0\}$, and $f : U \to \mathbb{R}$ be given by

$$f(x,y) = \arg(x+iy) = \begin{cases} \arccos \dfrac{x}{\sqrt{x^2+y^2}} & \text{if } y > 0, \\ \pi & \text{if } y = 0, \\ 2\pi - \arccos \dfrac{x}{\sqrt{x^2+y^2}} & \text{if } y < 0. \end{cases}$$

Then

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

Since $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ are both continuous on U, f is differentiable on U.

Definition 5.43. Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}^m$ be differentiable on U. f is said to be ***continuously differentiable*** on U if $Df : U \to \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$ is continuous on U. The collection of all continuously differentiable mappings from U to $\mathbb{R}^m$ is denoted by

$\mathscr{C}^1(U;\mathbb{R}^m)$. The collection of all bounded differentiable functions from U to $\mathbb{R}^m$ whose derivative is continuous and bounded is denoted by $\mathscr{C}_b^1(U;\mathbb{R}^m)$. In other words,

$$\mathscr{C}^1(U;\mathbb{R}^m) = \big\{ f : U \to \mathbb{R}^m \text{ is differentiable on } U \,\big|\, Df : U \to \mathscr{B}(\mathbb{R}^n,\mathbb{R}^m) \text{ is continuous} \big\}$$

and

$$\mathscr{C}_b^1(U;\mathbb{R}^m) = \Big\{ f \in \mathscr{C}^1(U;\mathbb{R}^m) \,\Big|\, \sup_{x\in U} |f(x)| + \sup_{x\in U} \|Df(x)\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} < \infty \Big\}.$$

Theorem 5.44. *Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}^m$. Then $f \in \mathscr{C}^1(U;\mathbb{R}^m)$ if and only if the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ exist and are continuous on U for $i = 1, \cdots, m$ and $j = 1, \cdots, n$.*

Proof. Note that $\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)$ is finite dimensional. By Example 4.29, there exist c and $C > 0$ such that

$$c\sum_{i=1}^m\sum_{j=1}^n |a_{ij}| \leqslant \|L\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} \leqslant C\sum_{i=1}^m\sum_{j=1}^n |a_{ij}| \quad \forall\, L \in \mathscr{B}(\mathbb{R}^n,\mathbb{R}^m) \text{ with representation } [L] = [a_{ij}].$$

Therefore, for every $a \in U$,

$$c\sum_{i=1}^m\sum_{j=1}^n \Big|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\Big| \leqslant \big\|(Df)(x) - (Df)(a)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} \leqslant C\sum_{i=1}^m\sum_{j=1}^n \Big|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\Big|;$$

thus $\lim\limits_{x\to a} \|(Df)(x) - (Df)(a)\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} = 0$ if and only if $\lim\limits_{x\to a} \Big|\dfrac{\partial f_i}{\partial x_j}(x) - \dfrac{\partial f_i}{\partial x_j}(a)\Big| = 0$ for all $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$. □

Example 5.45. If $f : \mathbb{R} \to \mathbb{R}$ is differentiable at a, must f' be continuous at a? In other words, is it always true that $\lim\limits_{x\to a} f'(x) = f'(a)$?

Answer: No! For example, take

$$f(x) = \begin{cases} x^2 \sin\dfrac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then f is differentiable at $x = 0$ since the limit

$$\lim_{h\to 0} \frac{f(0+h) - f(0)}{h} = \lim_{h\to 0} \frac{h^2 \sin\frac{1}{h}}{h} = \lim_{h\to 0} h \sin\frac{1}{h} = 0$$

exists. Therefore,

$$f'(x) = \begin{cases} 2x \sin\dfrac{1}{x} - \cos\dfrac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

However, $\lim\limits_{x\to 0} f'(x)$ does not exist.

5.5 The Product Rule and the Chain Rule

Theorem 5.46. *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed vector spaces over field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, $U \subseteq X$ be open, and $f : U \to Y$ and $g : U \to \mathbb{F}$ be differentiable at $a \in U$. Then $gf : U \to Y$ is differentiable at a, and*

$$D(gf)(a)(v) = g(a)(Df)(a)(v) + (Dg)(a)(v)f(a). \tag{5.5.1}$$

Moreover, if $g(a) \neq 0$, then $\dfrac{f}{g} : U \to Y$ is also differentiable at a, and $D(\dfrac{f}{g})(a) : X \to Y$ is given by

$$D\big(\frac{f}{g}\big)(a)(v) = \frac{g(a)\big((Df)(a)(v)\big) - (Dg)(a)(v)f(a)}{g^2(a)}. \tag{5.5.2}$$

Proof. We only prove (5.5.1), and (5.5.2) is left as an exercise.

Define $A : X \to Y$ by $A(v) = g(a)(Df)(a)(v) + (Dg)(a)(v)f(a)$. Then clearly $A \in \mathscr{L}(X, Y)$. Moreover,

$$\begin{aligned}
\|Av\|_Y &\leqslant \big\|g(a)(Df)(a)(v)\big\|_Y + \big\|(Dg)(a)(v)f(a)\big\|_Y \\
&\leqslant \big|g(a)\big|\big\|(Df)(a)(v)\big\|_Y + \big|(Dg)(a)(v)\big|\big\|f(a)\big\|_Y \\
&\leqslant \big|g(a)\big|\big\|(Df)(a)\big\|_{\mathscr{B}(X,Y)}\|v\|_X + \big\|(Dg)(a)\big\|_{\mathscr{B}(X,\mathbb{F})}\|v\|_X\|f(a)\|_Y \\
&= \Big[\big|g(a)\big|\big\|(Df)(a)\big\|_{\mathscr{B}(X,Y)} + \big\|(Dg)(a)\big\|_{\mathscr{B}(X,\mathbb{F})}\|f(a)\|_Y\Big]\|v\|_X
\end{aligned}$$

so that A is bounded. Note that

$$\begin{aligned}
(gf)(x) - (gf)(a) - A(x-a) = g(x)\big(&f(x) - f(a) - (Df)(a)(x-a)\big) \\
&+ \big(g(x) - g(a) - (Dg)(a)(x-a)\big)f(a) \\
&+ \big(g(x) - g(a)\big((Df)(a)(x-a)\big)\big).
\end{aligned}$$

As a consequence, for $x \neq a$,

$$\begin{aligned}
\frac{\big\|(gf)(x) - (gf)(a) - A(x-a)\big\|_Y}{\|x-a\|_X} &\leqslant \big|g(x)\big|\frac{\big\|f(x) - f(a) - (Df)(a)(x-a)\big\|_Y}{\|x-a\|_X} \\
&+ \frac{\big|g(x) - g(a) - (Dg)(a)(x-a)\big|}{\|x-a\|_X}\|f(a)\|_Y + \big\|(Df)(a)\big\|_{\mathscr{B}(X,Y)}\big|g(x) - g(a)\big|
\end{aligned}$$

and the right-hand side approaches zero as $x \to a$ since f and g are differentiable at a (so that g is continuous at a). Therefore, gf is differentiable at a with derivative $D(gf)(a)$ given by (5.5.1). □

Theorem 5.47. *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ be normed vector spaces, $U \subseteq X$ and $V \subseteq Y$ be open sets. Suppose that $f : U \to Y$ is differentiable at $a \in U$, $f(U) \subseteq V$, and $g : V \to Z$ is differentiable at $f(a)$. Then the map $F = g \circ f : U \to Z$ defined by*

$$F(x) = g\big(f(x)\big) \qquad \forall\, x \in U$$

is differentiable at a, and

$$(DF)(a)(h) = (Dg)\big(f(a)\big)\big((Df)(a)(h)\big) \qquad \forall\, h \in X\,.$$

In particular, if $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ and $Z = \mathbb{R}^\ell$, then

$$\big((DF)(a)\big)_{ij} = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}\big(f(a)\big)\frac{\partial f_k}{\partial x_j}(a)\,.$$

Proof. To simplify the notation, we write $b = f(a)$, $A = (Df)(a) \in \mathscr{B}(X, Y)$, and $B = (Dg)(b) \in \mathscr{B}(Y, Z)$. Since U and V are open, there exists $r_1, r_2 > 0$ such that $B_X(a, r_1) \subseteq U$ and $B_Y(b, r_2) \subseteq V$, where B_X and B_Y denote balls in X and Y, respectively.

Let $\varepsilon > 0$ be given. Define $u : B_X(0, r_1) \to Y$ and $v : B_Y(0, r_2) \to Z$ by

$$u(h) = f(a+h) - f(a) - Ah \qquad \text{and} \qquad v(k) = g(b+k) - g(b) - Bk\,.$$

By the differentiability of f and g at a and b, there exist $0 < \delta_1 < r_1$ and $0 < \delta_2 < r_2$ such that

$$\|u(h)\|_Y \leqslant \min\Big\{1, \frac{\varepsilon}{2(\|B\|+1)}\Big\}\|h\|_X \quad \text{whenever} \quad \|h\|_X < \delta_1\,,$$
$$\|v(k)\|_Z \leqslant \frac{\varepsilon}{2(\|A\|+1)}\|k\|_Y \quad \text{whenever} \quad \|k\|_Y < \delta_2\,.$$

Let $k = f(a+h) - f(a) = Ah + u(h)$. Then $\lim\limits_{h\to 0} k = 0$; thus there exists $\delta_3 > 0$ such that

$$\|k\|_Y < \delta_2 \quad \text{whenever} \quad \|h\|_X < \delta_3\,.$$

Define $\delta = \min\{\delta_1, \delta_3\}$. Then $\delta > 0$; thus by the fact that

$$\begin{aligned} F(a+h) - F(a) &= g(b+k) - g(b) = Bk + v(k) = B\big(Ah + u(h)\big) + v(k) \\ &= BAh + Bu(h) + v(k)\,, \end{aligned}$$

we find that if $\|h\|_X < \delta$,

$$\begin{aligned} \|F(a+h) - F(a) - BAh\|_Z &\leqslant \|Bu(h)\|_Z + \|v(k)\|_Z \leqslant \|B\|\|u(h)\|_Y + \frac{\varepsilon}{2(\|A\|+1)}\|k\|_Y \\ &\leqslant \frac{\varepsilon}{2}\|h\|_X + \frac{\varepsilon}{2(\|A\|+1)}\big(\|A\|\|h\|_X + \|u(h)\|_Y\big) \leqslant \frac{\varepsilon}{2}\|h\|_X + \frac{\varepsilon}{2}\|h\|_X = \varepsilon\|h\|_X\,. \end{aligned}$$

Since $BA = B \circ A \in \mathscr{B}(X,Z)$ by Proposition 5.9, we conclude that F is differentiable at a and $(DF)(a) = BA$. □

Example 5.48. Let $X = \mathscr{C}([a,b];\mathbb{R})$ and $\|\cdot\|_X$ be the maximum norm; that is, $\|f\|_X = \max\limits_{x\in[a,b]} \big|f(x)\big|$. Let $I : X \to X$ be defined by $I(f) = f^2$ and $J : X \to \mathbb{R}$ be defined by $J(f) = \displaystyle\int_a^b f(x)^2\,dx$. Then I is differentiable on X (with $(DI)(f)(h) = 2fh$) and J is differentiable on X $\big($with $(DJ)(f)(h) = \displaystyle\int_a^b 2f(x)h(x)\,dx\big)$. Therefore, the chain rule implies that $J \circ I$ is differentiable on X and

$$D(J\circ I)(f)(h) = (DJ)(f^2)\big((DI)(f)(h)\big) = (DJ)(f^2)(2fh) = \int_a^b 4f^3(x)h(x)\,dx\,.$$

Example 5.49. Let $f : \mathrm{GL}(n) \to \mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)$ and $g : \mathscr{B}(\mathbb{R}^n,\mathbb{R}^n) \to \mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)$ be defined by $f(L) = L^{-1}$ and $g(L) = L^2$. Then for $H \in \mathrm{GL}(n)$, Example 5.19 and 5.20 imply that

$$(Df)(L)(H) = -L^{-1}HL^{-1} \qquad \text{and} \qquad (Dg)(L)(H) = LH + HL\,.$$

Therefore, the chain rule shows that

$$\begin{aligned} D(g\circ f)(L)(H) &= (Dg)(L^{-1})\big((Df)(L)(H)\big) = L^{-1}\big((Df)(L)(H)\big) + \big((Df)(L)(H)\big)L^{-1} \\ &= -L^{-2}HL^{-1} - L^{-1}HL^{-2}\,. \end{aligned}$$

5.6 Higher Derivatives of Functions

Let $U \subseteq X$ be open, and $f : U \to Y$ is differentiable. By Proposition 5.8, the space $\big(\mathscr{B}(X,Y), \|\cdot\|_{\mathscr{B}(X,Y)}\big)$ is a normed space, so it is legitimate to ask if $Df : U \to \mathscr{B}(X,Y)$ is differentiable or not. If Df is differentiable at a, we call f twice differentiable at a, and denote the twice derivative of f at a as $(D^2f)(a)$. If Df is differentiable on U, then $D^2f : U \to \mathscr{B}\big(X, \mathscr{B}(X,Y)\big)$. Similar, we can talk about three times differentiability of a function if it is twice differentiable. In general, we have the following

Definition 5.50. Let $(X,\|\cdot\|_X)$ and $(Y,\|\cdot\|_Y)$ be normed spaces, and $U \subseteq X$ be open. A function $f : U \to Y$ is said to be ***twice differentiable*** at $a \in U$ if

1. f is (once) differentiable in a neighborhood of a;

2. there exists $L_2 \in \mathscr{B}\big(X, \mathscr{B}(X,Y)\big)$, usually denoted by $(D^2 f)(a)$ and called the ***second derivative*** of f at a, such that
$$\lim_{x\to a} \frac{\big\|(Df)(x) - (Df)(a) - L_2(x-a)\big\|_{\mathscr{B}(X,Y)}}{\|x-a\|_X} = 0\,.$$

For two vectors $u, v \in X$, $(D^2 f)(a)(v) \in \mathscr{B}(X,Y)$ and $(D^2 f)(a)(v)(u) \in Y$. The vector $(D^2 f)(a)(v)(u)$ is usually denoted by $(D^2 f)(a)(u,v)$.

In general, a function f is said to be ***k-times differentiable*** at $a \in U$ if

1. f is $(k-1)$-times differentiable in a neighborhood of a;
2. there exists $L_k \in \mathscr{B}(\underbrace{X, \mathscr{B}(X, \cdots, \mathscr{B}(X}_{k \text{ copies of "}X\text{"}}, Y\underbrace{)\cdots))}_{k \text{ copies of ")"}}$, usually denoted by $(D^k f)(a)$ and called the ***k-th derivative*** of f at a, such that
$$\lim_{x\to a} \frac{\big\|(D^{k-1} f)(x) - (D^{k-1} f)(a) - L_k(x-a)\big\|_{\mathscr{B}(X,\mathscr{B}(X,\cdots,\mathscr{B}(X,Y)\cdots))}}{\|x-a\|_X} = 0\,.$$

For k vectors $u^{(1)}, \cdots, u^{(k)} \in X$, the vector $(D^k f)(a)(u^{(1)}, \cdots, u^{(k)})$ is defined as the vector
$$(D^k f)(a)(u^{(k)})(u^{(k-1)}) \cdots (u^{(1)})\,,$$
where $(D^k f)(a)(u^{(k)}) \in \mathscr{B}(\underbrace{X, \mathscr{B}(X, \cdots, \mathscr{B}(X}_{(k-1) \text{ copies of "}X\text{"}}, Y\underbrace{)\cdots))}_{(k-1) \text{ copies of ")"}}$ so that $(D^k f)(a)(u^{(k)})(u^{(k-1)}) \in \mathscr{B}(\underbrace{X, \mathscr{B}(X, \cdots, \mathscr{B}(X}_{(k-2) \text{ copies of "}X\text{"}}, Y\underbrace{)\cdots))}_{(k-2) \text{ copies of ")"}}$, and etc.

Example 5.51. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed spaces, and $f(x) = Lx$ for some $L \in \mathscr{B}(X,Y)$. From Example 5.18, $(Df)(a) = L$ for all $a \in X$; thus $(D^2 f)(a) = 0$ since $Df : U \in \mathscr{B}(X,Y)$ is a "constant" map. In fact, one can also conclude from
$$\lim_{x\to a} \frac{\big\|(Df)(x) - (Df)(a) - 0(x-a)\big\|_{\mathscr{B}(X,Y)}}{\|x-a\|_X} = 0$$
that $(D^2 f)(a) = 0$ for all $a \in X$.

Remark 5.52. We focus on what $(D^k f)(a)(u_k)(\cdots)(u_1)$ means in this remark. We first look at the case that f is twice differentiable at a. With $x = a + tv$ for $v \in X$ with $\|v\|_X = 1$ in the definition, we find that
$$\lim_{t\to 0} \frac{\big\|(Df)(a+tv) - (Df)(a) - t(D^2 f)(a)(v)\big\|_{\mathscr{B}(X,Y)}}{|t|} = 0\,.$$

Since $(Df)(a+tv)-(Df)(a)-t(D^2f)(a)(v) \in \mathscr{B}(X,Y)$, for all $u \in X$ with $\|u\|_X = 1$ we have

$$\begin{aligned}
&\lim_{t\to 0}\Big\|\frac{(Df)(a+tv)(u)-(Df)(a)(u)}{t}-(D^2f)(a)(v)(u)\Big\|_Y \\
&\qquad = \lim_{t\to 0}\frac{\|(Df)(a+tv)(u)-(Df)(a)(u)-t(D^2f)(a)(v)(u)\|_Y}{|t|} \\
&\qquad = \lim_{t\to 0}\frac{\big\|\big[(Df)(a+tv)-(Df)(a)-t(D^2f)(a)(v)\big](u)\big\|_Y}{|t|} \\
&\qquad \leqslant \lim_{t\to 0}\frac{\|(Df)(a+tv)-(Df)(a)-t(D^2f)(a)(v)\|_{\mathscr{B}(X,Y)}}{\|(a+tv)-a\|_X} = 0\,.
\end{aligned}$$

On the other hand, by the definition of the direction derivative (see Remark 5.15),

$$(Df)(a+tv)(u)-(Df)(a)(u) = \lim_{s\to 0}\Big[\frac{f(a+tv+su)-f(a+tv)}{s}-\frac{f(a+su)-f(a)}{s}\Big];$$

thus the limit above implies that

$$\begin{aligned}
(D^2f)(a)(v)(u) &= \lim_{t\to 0}\lim_{s\to 0}\frac{f(a+tv+su)-f(a+tv)-f(a+su)+f(a)}{st} \\
&= \lim_{t\to 0}\frac{\displaystyle\lim_{s\to 0}\frac{f(a+tv+su)-f(a+tv)}{s}-\lim_{s\to 0}\frac{f(a+su)-f(a)}{s}}{t} \\
&= D_v(D_uf)(a)\,.
\end{aligned}$$

Therefore, $(D^2f)(a)(v)(u)$ is obtained by first differentiating f around a in the u-direction, then differentiating (Df) at a in the v-direction.

In general, $(D^kf)(a)(u_k)\cdots(u_1)$ is obtained by first differentiating f around a in the u_1-direction, then differentiating (Df) near a in the u_2-direction, and so on, and finally differentiating $(D^{k-1}f)$ at a in the u_k-direction.

Remark 5.53. Since $(D^2f)(a) \in \mathscr{B}(X,\mathscr{B}(X,Y))$, if $v_1,v_2 \in X$ and $c \in \mathbb{R}$, we have $(D^2f)(a)(cv_1+v_2) = c(D^2f)(a)(v_1)+(D^2f)(a)(v_2)$ (treated as "vectors" in $\mathscr{B}(X,Y)$); thus

$$(D^2f)(a)(cv_1+v_2)(u) = c(D^2f)(a)(v_1)(u)+(D^2f)(a)(v_2)(u) \qquad \forall\, u,v_1,v_2 \in X\,.$$

On the other hand, since $(D^2f)(a)(v) \in \mathscr{B}(X,Y)$,

$$(D^2f)(a)(v)(cu_1+u_2) = c(D^2f)(a)(v)(u_1)+(D^2f)(a)(v)(u_2) \qquad \forall\, u_1,u_2,v \in X\,.$$

Therefore, $(D^2f)(a)(v)(u)$ is linear in both u and v variables. A map with such kind of property is called a ***bilinear*** map (meaning 2-linear). In particular, $(D^2f)(a) : X\times X \to Y$ is a bilinear map.

In general, the vector $(D^k f)(a)(u^{(1)}, \cdots, u^{(k)})$ is linear in $u^{(1)}$, $\cdots$, $u^{(k)}$; that is,

$$\begin{aligned}(D^k f)(a)&(u^{(1)}, \cdots, u^{(i-1)}, \alpha v + \beta w, u^{(i+1)}, \cdots, u^{(k)})\\ &= \alpha(D^k f)(a)(u^{(1)}, \cdots, u^{(i-1)}, v, u^{(i+1)}, \cdots, u^{(k)})\\ &\quad + \beta(D^k f)(a)(u^{(1)}, \cdots, u^{(i-1)}, w, u^{(i+1)}, \cdots, u^{(k)})\end{aligned}$$

for all $v, w \in X$, $\alpha, \beta \in \mathbb{R}$, and $i = 1, \cdots, n$. Such kind of map which is linear in each component when the other $k-1$ components are fixed is called ***k*-linear**.

Consider the case that X is finite dimensional with $\dim(X) = n$, $\{\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n\}$ is a basis of X, and $Y = \mathbb{R}$. Then $(D^2 f)(a) : X \times X \to Y$ is a bilinear form (here the term "form" means that $Y = \mathbb{R}$). A bilinear form $B : X \times X \to \mathbb{R}$ can be represented as follows: Let $a_{ij} = B(\mathrm{e}_i, \mathrm{e}_j) \in \mathbb{R}$ for $i, j = 1, 2, \cdots, n$. Given $x, y \in \mathbb{R}^n$, write $u = \sum\limits_{i=1}^{n} u_i \mathrm{e}_i$ and $v = \sum\limits_{j=1}^{n} v_j \mathrm{e}_j$. Then by the bilinearity of B,

$$B(u, v) = B\Big(\sum_{i=1}^{n} u_i \mathrm{e}_i, \sum_{j=1}^{n} v_j \mathrm{e}_j\Big) = \sum_{i,j=1}^{n} u_i v_j a_{ij} = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

Therefore, if $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is twice differentiable at a, then the bilinear form $(D^2 f)(a)$ can be represented as

$$(D^2 f)(a)(u, v) = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} (D^2 f)(\mathrm{e}_1, \mathrm{e}_1) & \cdots & (D^2 f)(a)(\mathrm{e}_1, \mathrm{e}_n) \\ \vdots & \ddots & \vdots \\ (D^2 f)(\mathrm{e}_n, \mathrm{e}_1) & \cdots & (D^2 f)(a)(\mathrm{e}_n, \mathrm{e}_n) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

The following proposition is an analogy of Proposition 5.39. The proof is similar to the one of Proposition 5.39, and is left as an exercise.

Proposition 5.54. *Let $U \subseteq \mathbb{R}^n$ be open, $a \in U$, and $f = (f_1, \cdots, f_m) : U \to \mathbb{R}^m$. Then f is k-times differentiable at a if and only if f_i is k-times differentiable at a for all $i = 1, \cdots, m$.*

Due to the proposition above, when talking about the higher-order differentiability of $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and a point $a \in U$, from now on we only focus on the case $m = 1$.

Example 5.55. In this example, we focus on what the second derivative $(D^2 f)(a)$ of a function f is, or in particular, what $(D^2 f)(a)(\mathrm{e}_i, \mathrm{e}_j)$ (which appears in the Remark 5.53) is, if $X = \mathbb{R}^2$.

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be differentiable, then

$$\big[(Df)(x,y)\big] = \big[f_x(x,y) \quad f_y(x,y)\big] = \left[\frac{\partial f}{\partial x}(x,y) \quad \frac{\partial f}{\partial y}(x,y)\right].$$

Suppose that f is twice differentiable at (a,b), and let $L_2 = (D^2 f)(a,b)$. Then

$$\lim_{(x,y)\to(a,b)} \frac{\big\|(Df)(x,y) - (Df)(a,b) - L_2\big((x-a, y-b)\big)\big\|_{\mathscr{B}(\mathbb{R}^2,\mathbb{R})}}{\sqrt{(x-a)^2 + (y-b)^2}} = 0$$

or equivalently,

$$\lim_{(x,y)\to(a,b)} \frac{\big\|\, \big[f_x(x,y) \quad f_y(x,y)\big] - \big[f_x(a,b) \quad f_y(a,b)\big] - \big[L_2\big((x-a, y-b)\big)\big]\big\|_{\mathscr{B}(\mathbb{R}^2,\mathbb{R})}}{\sqrt{(x-a)^2 + (y-b)^2}} = 0\,,$$

where $\big[L_2\big((x-a, y-b)\big)\big]$ denotes the matrix representation of the linear map $L_2\big((x-a, y-b)\big) \in \mathscr{B}(\mathbb{R}^2,\mathbb{R})$. In particular, we must have

$$\lim_{x\to a} \left\| \left[\frac{f_x(x,b) - f_x(a,b)}{x-a} \quad \frac{f_y(x,b) - f_y(a,b)}{x-a}\right] - \big[L_2 \mathrm{e}_1\big] \right\|_{\mathscr{B}(\mathbb{R}^2,\mathbb{R})} = 0$$

and

$$\lim_{y\to b} \left\| \left[\frac{f_x(a,y) - f_x(a,b)}{y-b} \quad \frac{f_y(a,y) - f_y(a,b)}{y-b}\right] - \big[L_2 \mathrm{e}_2\big] \right\|_{\mathscr{B}(\mathbb{R}^2,\mathbb{R})} = 0\,.$$

Using the notation of second partial derivatives, we find that

$$\big[L_2 \mathrm{e}_1\big] = \big[f_{xx}(a,b) \quad f_{yx}(a,b)\big] \quad \text{and} \quad \big[L_2 \mathrm{e}_2\big] = \big[f_{xy}(a,b) \quad f_{yy}(a,b)\big]\,,$$

where $f_{xy} = (f_x)_y = \dfrac{\partial}{\partial y}\Big(\dfrac{\partial f}{\partial x}\Big)$ and $f_{yx} = (f_y)_x = \dfrac{\partial}{\partial x}\Big(\dfrac{\partial f}{\partial y}\Big)$. Therefore, if $v = v_1 \mathrm{e}_1 + v_2 \mathrm{e}_2$,

$$[L_2 v] = \big[L_2(v_1 \mathrm{e}_1 + v_2 \mathrm{e}_2)\big] = \big[v_1 f_{xx}(a,b) + v_2 f_{xy}(a,b) \quad v_1 f_{yx}(a,b) + v_2 f_{yy}(a,b)\big]\,. \tag{5.6.1}$$

Symbolically, we can write

$$\big[L_2\big] = \Big[\, \big[f_{xx}(a,b) \quad f_{yx}(a,b)\big] \quad \big[f_{xy}(a,b) \quad f_{yy}(a,b)\big] \,\Big]$$

so that

$$\begin{aligned}\big[L_2(v_1 \mathrm{e}_1 + v_2 \mathrm{e}_2)\big] &= \big[L_2\big] \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \Big[\, \big[f_{xx}(a,b) \quad f_{yx}(a,b)\big] \quad \big[f_{xy}(a,b) \quad f_{yy}(a,b)\big] \,\Big] \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= v_1 \big[f_{xx}(a,b) \quad f_{yx}(a,b)\big] + v_2 \big[f_{xy}(a,b) \quad f_{yy}(a,b)\big]\,.\end{aligned}$$

For two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, what does $(D^2f)(a,b)(\boldsymbol{v})(\boldsymbol{u})$ or $(D^2f)(a,b)(\boldsymbol{u},\boldsymbol{v})$ mean? To see this, let $\boldsymbol{u} = u_1\mathrm{e}_1 + u_2\mathrm{e}_2$ and $\boldsymbol{v} = v_1\mathrm{e}_1 + v_2\mathrm{e}_2$. Then

$$\begin{aligned}
\big[(D^2f)(a,b)(\boldsymbol{v})(\boldsymbol{u})\big] &= \big[(D^2f)(a,b)(\boldsymbol{v})\big]\big[\boldsymbol{u}\big] = \big[L_2(v_1\mathrm{e}_1 + v_2\mathrm{e}_2)\big]\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\
&= v_1\begin{bmatrix} f_{xx}(a,b) & f_{yx}(a,b) \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + v_2\begin{bmatrix} f_{xy}(a,b) & f_{yy}(a,b) \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\
&= \begin{bmatrix} v_1 & v_2 \end{bmatrix}\begin{bmatrix} f_{xx}(a,b) & f_{yx}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.
\end{aligned}$$

Therefore, $(D^2f)(a,b)(\mathrm{e}_1,\mathrm{e}_1) = f_{xx}(a,b)$, $(D^2f)(a,b)(\mathrm{e}_1,\mathrm{e}_2) = f_{xy}(a,b)$, $(D^2f)(a,b)(\mathrm{e}_2,\mathrm{e}_1) = f_{yx}(a,b)$ and $(D^2f)(a,b)(\mathrm{e}_2,\mathrm{e}_2) = f_{yy}(a,b)$.

On the other hand, we can identify $\mathscr{B}(\mathbb{R}^2;\mathbb{R})$ as $\mathbb{R}^2$ (every 1×2 matrix is a "row" vector), and treat $g \equiv \big[Df\big]^{\mathrm{T}} : \mathbb{R}^2 \to \mathbb{R}^2$ as a vector-valued function. By Theorem 5.27 $(Dg)(a,b)$ can be represented as a 2×2 matrix given by

$$\big[(Dg)(a,b)\big] = \begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{bmatrix}.$$

We note that the representation above means

$$\lim_{(x,y)\to(a,b)} \frac{\left\| \begin{bmatrix} f_x(x,y) \\ f_y(x,y) \end{bmatrix} - \begin{bmatrix} f_x(a,b) \\ f_y(a,b) \end{bmatrix} - \begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{bmatrix}\begin{bmatrix} x-a \\ y-b \end{bmatrix} \right\|_{\mathbb{R}^2}}{\sqrt{(x-a)^2+(y-b)^2}} = 0.$$

The equality above is equivalent to that

$$\lim_{(x,y)\to(a,b)} \frac{\left\| \big[(Df)(x,y)\big] - \big[(Df)(a,b)\big] - \begin{bmatrix} x-a & y-b \end{bmatrix}\begin{bmatrix} f_{xx}(a,b) & f_{yx}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{bmatrix} \right\|_{\mathbb{R}^2}}{\sqrt{(x-a)^2+(y-b)^2}} = 0$$

According to the equality above, $L_2 = (D^2f)(a,b)$ should be defined by

$$\big[L_2(v_1\mathrm{e}_1 + v_2\mathrm{e}_2)\big] = \begin{bmatrix} v_1 & v_2 \end{bmatrix}\begin{bmatrix} f_{xx}(a,b) & f_{yx}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{bmatrix} = \left(\begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right)^{\mathrm{T}}$$

which agrees with what (5.6.1) provides.

Proposition 5.56. *Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$. Suppose that f is k-times differentiable at a. Then for k vectors $u^{(1)}, \cdots, u^{(k)} \in \mathbb{R}^n$,*

$$(D^kf)(a)(u^{(1)}, \cdots, u^{(k)}) = \sum_{j_1,\cdots,j_k=1}^{n} \frac{\partial^k f}{\partial x_{j_k}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a)u^{(1)}_{j_1}u^{(2)}_{j_2}\cdots u^{(k)}_{j_k},$$

where $u^{(i)} = (u_1^{(i)}, u_2^{(i)}, \cdots, u_n^{(i)})$ *for all* $i = 1, \cdots, k$（上標括號中的數字指所給定的 k 個向量中的第幾個向量，下標指每一個固定向量的第幾個分量）*and*

$$\frac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} \cdots \partial x_{j_1}}(a) = \frac{\partial}{\partial x_{j_k}}\Big|_{x=a}\Big(\frac{\partial}{\partial x_{j_{k-1}}}\Big(\cdots \frac{\partial}{\partial x_{j_2}}\Big(\frac{\partial f}{\partial x_{j_1}}\Big)\cdots\Big)\Big).$$

Proof. We prove the proposition by induction. Let $\{\mathrm{e}_j\}_{j=1}^n$ be the standard basis of $\mathbb{R}^n$. By Remark 5.53 (on multi-linearity), it suffices to show that

$$(D^k f)(a)(\mathrm{e}_{j_k})(\mathrm{e}_{j_{k-1}})\cdots(\mathrm{e}_{j_2})(\mathrm{e}_{j_1}) = (D^k f)(a)(\mathrm{e}_{j_1}, \cdots, \mathrm{e}_{j_k}) = \frac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} \cdots \partial x_{j_1}}(a) \tag{5.6.2}$$

provided that f is k-times differentiable at a since if so, we must have

$$\begin{aligned}(D^k f)(a)(u^{(1)}, \cdots, u^{(k)}) &= (D^k f)(a)\Big(\sum_{j_1=1}^n u_{j_1}^{(1)} \mathrm{e}_{j_1}, \cdots, \sum_{j_k=1}^n u_{j_k}^{(k)} \mathrm{e}_{j_k}\Big) \\ &= \sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_k=1}^n (D^k f)(a)(\mathrm{e}_{j_1}, \cdots, \mathrm{e}_{j_k}) u_{j_1}^{(1)} u_{j_2}^{(2)} \cdots u_{j_k}^{(k)} \\ &= \sum_{j_1, \cdots, j_k=1}^n \frac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} \cdots \partial x_{j_1}}(a) u_{j_1}^{(1)} u_{j_2}^{(2)} \cdots u_{j_k}^{(k)}.\end{aligned}$$

Note that the case $k = 1$ is true because of Theorem 5.27. Next we assume that (5.6.2) holds true for $k = \ell$ if f is $(\ell - 1)$-times differentiable in a neighborhood of a and f is ℓ-times differentiable at a. Now we show that (5.6.2) also holds true for $k = \ell + 1$ if f is ℓ-times differentiable in a neighborhood of a, and f is $(\ell + 1)$-times differentiable at a. By the definition of $(\ell + 1)$-times differentiability at a,

$$\lim_{x \to a} \frac{\big\|(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x - a)\big\|_{\mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n, \cdots, \mathscr{B}(\mathbb{R}^n, \mathbb{R})\cdots))}}{\|x - a\|_{\mathbb{R}^n}} = 0.$$

Since

$$\begin{aligned}&\Big|\big[(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x - a)\big](\mathrm{e}_{j_\ell})\cdots(\mathrm{e}_{j_2})(\mathrm{e}_{j_1})\Big| \\ &\qquad \leqslant \Big\|\big[(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x - a)\big](\mathrm{e}_{j_\ell})\cdots(\mathrm{e}_{j_2})\Big\|_{\mathscr{B}(\mathbb{R}^n, \mathbb{R})} \|\mathrm{e}_{j_1}\|_{\mathbb{R}^n} \\ &\qquad \leqslant \big\|(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x - a)\big\|_{\mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n, \cdots, \mathscr{B}(\mathbb{R}^n, \mathbb{R})\cdots))} \|\mathrm{e}_{j_1}\|_{\mathbb{R}^n} \cdots \|\mathrm{e}_{j_\ell}\|_{\mathbb{R}^n} \\ &\qquad = \big\|(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x - a)\big\|_{\mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n, \cdots, \mathscr{B}(\mathbb{R}^n, \mathbb{R})\cdots))},\end{aligned}$$

using (5.6.2) (for the case $k = \ell$) we conclude that for $x \neq a$,

$$\frac{\left|\dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(x) - \dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a) - (D^{\ell+1}f)(a)(\mathrm{e}_{j_1},\cdots,\mathrm{e}_{j_\ell},x-a)\right|}{\|x-a\|_{\mathbb{R}^n}}$$
$$= \frac{\left|(D^\ell f)(x)(\mathrm{e}_{j_1},\cdots,\mathrm{e}_{j_\ell}) - (D^\ell f)(a)(\mathrm{e}_{j_1},\cdots,\mathrm{e}_{j_\ell}) - (D^{\ell+1}f)(a)(x-a)(\mathrm{e}_{j_1},\cdots,\mathrm{e}_{j_\ell})\right|}{\|x-a\|_{\mathbb{R}^n}}$$
$$\leqslant \frac{\left\|(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1}f)(a)(x-a)\right\|_{\mathscr{B}(\mathbb{R}^n,\mathscr{B}(\mathbb{R}^n,\cdots,\mathscr{B}(\mathbb{R}^n,\mathbb{R})\cdots))}}{\|x-a\|_{\mathbb{R}^n}}$$

and the right-hand side approaches zero as $x \to a$ so that

$$\lim_{x\to a}\frac{\left|\dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(x) - \dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a) - (D^{\ell+1}f)(a)(\mathrm{e}_{j_1},\cdots,\mathrm{e}_{j_\ell},x-a)\right|}{\|x-a\|_{\mathbb{R}^n}} = 0\,.$$

In particular, we pass to the limit as $x \to a$ in the way $x = a + t\mathrm{e}_{j_{\ell+1}}$ as $t \to 0$ for some $j_{\ell+1} = 1,\cdots,n$ and conclude from the definition of partial derivatives that

$$\begin{aligned}(D^{\ell+1}f)(a)(\mathrm{e}_{j_1},\cdots,\mathrm{e}_{j_\ell},\mathrm{e}_{j_{\ell+1}}) &= \lim_{t\to 0}\frac{\dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a+t\mathrm{e}_{j_{\ell+1}}) - \dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a)}{t}\\ &= \frac{\partial^{\ell+1} f}{\partial x_{j_{\ell+1}}\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a)\end{aligned}$$

which is (5.6.2) for the case $k = \ell + 1$. $\square$

Example 5.57. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x_1, x_2) = x_1^2 \cos x_2$, and $u^{(1)} = (2, 0)$, $u^{(2)} = (1, 1)$, $u^{(3)} = (0, -1)$. Suppose that f is three-times differentiable at $a = (0, 0)$ (in fact it is, but we have not talked about this yet). Then

$$\begin{aligned}(D^3f)(a)(u^{(1)},u^{(2)},u^{(3)}) &= \sum_{i,j,k=1}^{2}\frac{\partial^3 f}{\partial x_k\partial x_j\partial x_i}(a)u_i^{(1)}u_j^{(2)}u_k^{(3)} = \sum_{j=1}^{2}\frac{\partial^3 f}{\partial x_2\partial x_j\partial x_1}(a)\cdot 2\cdot u_j^{(2)}\cdot(-1)\\ &= \frac{\partial^3 f}{\partial x_2\partial x_1^2}(0,0)\cdot 2\cdot 1\cdot(-1) + \frac{\partial^3 f}{\partial x_2^2\partial x_1}(0,0)\cdot 2\cdot 1\cdot(-1) = 0\,.\end{aligned}$$

Corollary 5.58. *Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$ be $(k+1)$-times differentiable at a. Then for $u^{(1)},\cdots,u^{(k)},u^{(k+1)} \in \mathbb{R}^n$,*

$$(D^{k+1}f)(a)(u^{(1)},\cdots,u^{(k)},u^{(k+1)}) = \sum_{j=1}^{n} u_j^{(k+1)}\frac{\partial}{\partial x_j}\Big|_{x=a}(D^kf)(x)(u^{(1)},\cdots,u^{(k)})\,.$$

In other words, (using the terminology in Remark 5.15*) $(D^{k+1}f)(a)(u^{(1)},\cdots,u^{(k)},u^{(k+1)})$ is the "directional derivative" of the function $(D^kf)(\cdot)(u^{(1)},\cdots,u^{(k)})$ at a in the "direction" $u^{(k+1)}$.*

Proof. By Proposition 5.56,

$$\begin{aligned}
(D^{k+1}f)(a)(u^{(1)},\cdots,u^{(k)},u^{(k+1)}) &= \sum_{j_1,\cdots,j_k,j_{k+1}=1}^{n} \frac{\partial^{k+1} f}{\partial x_{j_{k+1}}\partial x_{j_k}\cdots\partial x_{j_1}}(a)u^{(1)}_{j_1}\cdots u^{(k)}_{j_k}u^{(k+1)}_{j_{k+1}} \\
&= \sum_{j_{k+1}=1}^{n} u^{(k+1)}_{j_{k+1}} \sum_{j_1,\cdots,j_k=1}^{n} \frac{\partial^{k+1} f}{\partial x_{j_{k+1}}\partial x_{j_k}\cdots\partial x_{j_1}}(a)u^{(1)}_{j_1}\cdots u^{(k)}_{j_k} \\
&= \sum_{j_{k+1}=1}^{n} u^{(k+1)}_{j_{k+1}} \frac{\partial}{\partial x_{j_{k+1}}}\Big|_{x=a} \sum_{j_1,\cdots,j_k=1}^{n} \frac{\partial^{k} f}{\partial x_{j_k}\cdots\partial x_{j_1}}(x)u^{(1)}_{j_1}\cdots u^{(k)}_{j_k} \\
&= \sum_{j_{k+1}=1}^{n} u^{(k+1)}_{j_{k+1}} \frac{\partial}{\partial x_{j_{k+1}}}\Big|_{x=a} (D^k f)(x)(u^{(1)},\cdots,u^{(k)}).
\end{aligned}$$

□

Example 5.59. Let $f:\mathbb{R}^2\to\mathbb{R}$ be twice differentiable at $a=(a_1,a_2)\in\mathbb{R}^2$. Then the proposition above shows that for $u=(u_1,u_2), v=(v_1,v_2)\in\mathbb{R}^2$,

$$\begin{aligned}
(D^2f)(a)(v)(u) = (D^2f)(a)(u,v) &= \sum_{i,j=1}^{2} \frac{\partial^2 f}{\partial x_j\partial x_i}(a)u_iv_j \\
&= \frac{\partial^2 f}{\partial x_1^2}(a)u_1v_1 + \frac{\partial^2 f}{\partial x_2\partial x_1}(a)u_1v_2 + \frac{\partial^2 f}{\partial x_1\partial x_2}(a)u_2v_1 + \frac{\partial^2 f}{\partial x_2^2}(a)u_2v_2 \\
&= \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(a) & \dfrac{\partial^2 f}{\partial x_2\partial x_1}(a) \\ \dfrac{\partial^2 f}{\partial x_1\partial x_2}(a) & \dfrac{\partial^2 f}{\partial x_2^2}(a) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.
\end{aligned}$$

In general, if $f:\mathbb{R}^n\to\mathbb{R}$ be twice differentiable at $a=(a_1,\cdots,a_n)\in\mathbb{R}^n$. Then for $u=(u_1,\cdots,u_n), v=(v_1,\cdots,v_n)\in\mathbb{R}^n$,

$$(D^2f)(a)(v)(u) = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n\partial x_1}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_1\partial x_n}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(a) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

The bilinear form $B:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ given by

$$B(u,v) = (D^2f)(a)(v)(u) \qquad \forall\, u,v\in\mathbb{R}^n$$

is called the ***Hessian*** of f, and is represented (in the matrix form) as an $n \times n$ matrix by

$$\begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(a) \end{bmatrix}.$$

If the second partial derivatives $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)$ of f at a exists for all $i, j = 1, \cdots, n$ (here the twice differentiability of f at a is ignored), the matrix (on the right-hand side of equality) above is also called the ***Hessian matrix*** of f at a.

Even though there is no reason to believe that $(D^2f)(a)(u,v) = (D^2f)(a)(v,u)$ (since the left-hand side means first differentiating f in u-direction and then differentiating Df in v-direction, while the right-hand side means first differentiating f in v-direction then differentiating Df in u-direction), it is still reasonable to ask whether $(D^2f)(a)$ is symmetric or not; that is, could it be true that $(D^2f)(a)(u,v) = (D^2f)(a)(v,u)$ for all $u, v \in \mathbb{R}^n$? When f is twice differentiable at a, this is equivalent of asking (by plugging in $u = \mathrm{e}_i$ and $v = \mathrm{e}_j$) that whether or not

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a). \tag{5.6.3}$$

The following example provides a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that (5.6.3) does not hold at $a = (0,0)$. We remark that the function in the following example is not twice differentiable at a even though the Hessian matrix of f at a can still be computed.

Example 5.60. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Then

$$f_x(x,y) = \begin{cases} \dfrac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0), \end{cases}$$

and

$$f_y(x,y) = \begin{cases} \dfrac{x^5-4x^3y^2-xy^4}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0), \end{cases}$$

It is clear that f_x and f_y are continuous on $\mathbb{R}^2$; thus f is differentiable on $\mathbb{R}^2$. However,

$$f_{xy}(0,0) = \lim_{k\to 0} \frac{f_x(0,k) - f_x(0,0)}{k} = -1\,,$$

while

$$f_{yx}(0,0) = \lim_{h\to 0} \frac{f_y(h,0) - f_y(0,0)}{h} = 1\,;$$

thus the Hessian matrix of f at the origin is not symmetric.

Theorem 5.61 (Clairaut's Theorem)**.** *Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$. Suppose that the mixed partial derivatives $\dfrac{\partial f}{\partial x_i}$, $\dfrac{\partial f}{\partial x_j}$, $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$, $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ exist in a neighborhood of a, and $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ is continuous at a. Then*

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\,. \tag{5.6.4}$$

Proof. Let $a \in U$ be given. For real numbers $h, k \neq 0$ such that $a + h\mathrm{e}_i + k\mathrm{e}_j \in U$, define

$$Q(h,k) \equiv \frac{f(a + h\mathrm{e}_i + k\mathrm{e}_j) - f(a + h\mathrm{e}_i) - f(a + k\mathrm{e}_j) + f(a)}{hk}\,.$$

Then

$$\begin{aligned}\lim_{k\to 0} Q(h,k) &= \frac{1}{h} \lim_{k\to 0} \Big(\frac{f(a + h\mathrm{e}_i + k\mathrm{e}_j) - f(a + h\mathrm{e}_i)}{k} - \frac{f(a + k\mathrm{e}_j) + f(a)}{k} \Big) \\ &= \frac{1}{h} \Big(\frac{\partial f}{\partial x_j}(a + h\mathrm{e}_j) - \frac{\partial f}{\partial x_j}(a) \Big)\,;\end{aligned}$$

thus

$$\lim_{h\to 0} \lim_{k\to 0} Q(h,k) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\,. \tag{5.6.5}$$

Define $\varphi(x) = f(x + k\mathrm{e}_j) - f(x)$. Then the mean value theorem implies that

$$\begin{aligned}Q(h,k) &= \frac{\varphi(a + h\mathrm{e}_i) - \varphi(a)}{hk} = \frac{1}{k} \frac{\partial \varphi}{\partial x_i}(a + \theta_1 h\mathrm{e}_i) \\ &= \frac{1}{k} \Big(\frac{\partial f}{\partial x_i}(a + \theta_1 h\mathrm{e}_i + k\mathrm{e}_j) - \frac{\partial f}{\partial x_i}(a + \theta_1 h\mathrm{e}_i) \Big) \\ &= \frac{\partial^2 f}{\partial x_j \partial x_i}(a + \theta_1 h\mathrm{e}_i + \theta_2 k\mathrm{e}_j)\end{aligned}$$

for some functions $\theta_1 = \theta_1(h,k)$ and $\theta_2 = \theta_2(h,k)$ satisfying $0 < \theta_1, \theta_2 < 1$. Therefore, we establish that there exist functions $\theta_1 = \theta_1(h,k)$ and $\theta_2 = \theta_2(h,k)$ such that $\theta_1, \theta_2 \in (0,1)$ and

$$Q(h,k) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a + \theta_1 h\mathrm{e}_i + \theta_2 k\mathrm{e}_j).$$

Passing to the limit as $k \to 0$ first then $h \to 0$, using (5.6.5) and the continuity of $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ we conclude that

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \lim_{h\to 0}\lim_{k\to 0} Q(h,k) = \lim_{h\to 0}\lim_{k\to 0} \frac{\partial^2 f}{\partial x_j \partial x_i}(a + \theta_1 h\mathrm{e}_i + \theta_2 k\mathrm{e}_j) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a). \qquad \square$$

Remark 5.62. In view of Remark 5.52, (5.6.4) is the same as the following identity

$$\begin{aligned}
&\lim_{h\to 0}\lim_{k\to 0} \frac{f(a + h\mathrm{e}_i + k\mathrm{e}_j) - f(a + h\mathrm{e}_i) - f(a + k\mathrm{e}_j) + f(a)}{hk} \\
&\qquad = \lim_{k\to 0}\lim_{h\to 0} \frac{f(a + h\mathrm{e}_i + k\mathrm{e}_j) - f(a + h\mathrm{e}_i) - f(a + k\mathrm{e}_j) + f(a)}{hk}
\end{aligned}$$

which implies that the order of the two limits $\lim\limits_{h\to 0}$ and $\lim\limits_{k\to 0}$ can be interchanged without changing the value of the limit (under certain conditions).

Example 5.63. Let $f(x,y) = yx^2 \cos y^2$. Then

$$\begin{aligned}
f_{xy}(x,y) &= (2xy\cos y^2)_y = 2x\cos y^2 - 2xy(2y)\sin y^2 = 2x\cos y^2 - 4xy^2\sin y^2, \\
f_{yx}(x,y) &= (x^2\cos y^2 - yx^2(2y)\sin y^2)_x = (x^2\cos y^2 - 2x^2y^2\sin y^2)_x \\
&= 2x\cos y^2 - 4xy^2\sin y^2 = f_{xy}(x,y).
\end{aligned}$$

Definition 5.64. A function is said to be ***of class*** $\mathscr{C}^k$ if the first k derivatives exist and are continuous. A function is said to be ***smooth*** or ***of class*** $\mathscr{C}^\infty$ if it is of class $\mathscr{C}^k$ for all positive integer k.

Now we would like to answer the question of what kind of functions are k-times differentiable. Suppose that $U \subseteq \mathbb{R}^n$ is open and $f : U \to \mathbb{R}$. Note that by the definition of differentiability, f is k-times differentiable in U if and only if $D^{k-1}f$ is differentiable in U. This would further imply that f is k-times differentiable in U if and only if $D^{k-2}f$ is twice differentiable in U. Therefore, Proposition 5.39 and Theorem 5.44 imply that

$$\begin{aligned}
&f \text{ is } k\text{-times (continuously) differentiable in } U\\
&\quad\Leftrightarrow Df \text{ is } (k-1)\text{-times (continuously) differentiable in } U\\
&\quad\Leftrightarrow \Big[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n}\Big] \text{ is } (k-1)\text{-times (continuously) differentiable in } U\\
&\quad\Leftrightarrow \frac{\partial f}{\partial x_{j_1}} \text{ is } (k-1)\text{-times (continuously) differentiable in } U \text{ for all } 1 \leqslant j_1 \leqslant n\\
&\quad\Leftrightarrow D\frac{\partial f}{\partial x_{j_1}} \text{ is } (k-2)\text{-times (continuously) differentiable in } U \text{ for all } 1 \leqslant j_1 \leqslant n\\
&\quad\Leftrightarrow \Big[\frac{\partial^2 f}{\partial x_1 \partial x_{j_1}}, \cdots, \frac{\partial^2 f}{\partial x_n \partial x_{j_1}}\Big] \text{ is } (k-2)\text{-times (continuously) differentiable in } U\\
&\qquad\quad \text{for all } 1 \leqslant j_1 \leqslant n\\
&\quad\Leftrightarrow \frac{\partial^2 f}{\partial x_{j_2} \partial x_{j_1}} \text{ is } (k-2)\text{-times (continuously) differentiable in } U \text{ for all } 1 \leqslant j_1, j_2 \leqslant n.
\end{aligned}$$

Applying similar argument several times, we obtain the following theorem which is an analogy of Theorem 5.44.

Theorem 5.65. *Let $U \to \mathbb{R}^n$ and $f : U \to \mathbb{R}$. Suppose that the partial derivative $\dfrac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} \cdots \partial x_{j_1}}$ exists in a neighborhood of $a \in U$ and is continuous at a for all $j_1, \cdots, j_k = 1, \cdots, n$. Then f is k-times differentiable at a. Moreover, $\dfrac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} \cdots \partial x_{j_1}}$ is continuous on U if and only if f is of class $\mathscr{C}^k$.*

Corollary 5.66. *Let $U \subseteq \mathbb{R}^n$ be open, and f is of class $\mathscr{C}^2$. Then*

$$(D^2 f)(a)(u, v) = (D^2 f)(a)(v, u) \qquad \forall\, a \in U \text{ and } u, v \in \mathbb{R}^n.$$

5.7 Taylor's Theorem

Recall Taylor's Theorem for functions of one variable:

Theorem 5.67 (Taylor)**.** *Suppose that for some $k \in \mathbb{N}$, $f : (a, b) \to \mathbb{R}$ be $(k+1)$-times differentiable and $c \in (a, b)$. Then for all $x \in (a, b)$, there exists d in between c and x such that*

$$f(x) = \sum_{\ell=0}^{k} \frac{f^{(\ell)}(c)}{\ell!}(x-c)^\ell + \frac{f^{(k+1)}(d)}{(k+1)!}(x-c)^{(k+1)},$$

where $f^{(\ell)}$ denotes the ℓ-th derivative of f.

Theorem 5.68 (Taylor). *Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$ be $(k+1)$-times differentiable. Suppose that $x, a \in U$ and the line segment joining x and a lies in U. Then there exists a point c on the line segment joining x and a such that*

$$\begin{aligned} f(x) - f(a) = &\sum_{\ell=1}^{k} \frac{1}{\ell!}(D^\ell f)(a)(\overbrace{x-a, \cdots, x-a}^{\ell \text{ copies of } x-a}) \\ &+ \frac{1}{(k+1)!}(D^{k+1} f)(c)(\underbrace{x-a, \cdots, x-a}_{(k+1) \text{ copies of } x-a}). \end{aligned} \tag{5.7.1}$$

Proof. Let $g(t) = f\big((1-t)a + tx\big)$. Since $\overline{xa} \subseteq U$ and U is open, there exists $\delta > 0$ such that $(1-t)a + tx \in U$ for all $t \in (-\delta, 1+\delta)$. By the chain rule, for $t \in (-\delta, 1+\delta)$,

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\big((1-t)a + tx\big)(x_i - a_i) = (Df)\big((1-t)a + tx\big)(x-a);$$

thus for $t \in (-\delta, 1+\delta)$, Proposition 5.56 shows that

$$g''(t) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}\big((1-t)a + tx\big)(x_i - a_i)(x_j - a_j) = (D^2 f)\big((1-t)a + tx\big)(x-a, x-a).$$

By induction, we conclude that

$$g^{(\ell)}(t) = (D^\ell f)\big((1-t)a + tx\big)(\underbrace{x-a, \cdots, x-a}_{\ell \text{ copies of } x-a}).$$

By the fact that f is $(k+1)$-times differentiable, $g : (-\delta, 1+\delta) \to \mathbb{R}$ is $(k+1)$-times differentiable as well. Theorem 5.67 then implies that for some $t_0 \in (0,1)$,

$$g(1) - g(0) = \sum_{\ell=1}^{k} \frac{g^{(\ell)}(0)}{\ell!} + \frac{g^{(k+1)}(t_0)}{(k+1)!}. \tag{5.7.2}$$

Letting $c = (1-t_0)a + t_0 x$, (5.7.2) implies (5.7.1). □

Definition 5.69. Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$ be k-times differentiable. The k-th (order) Taylor polynomial for f at a is the polynomial

$$\sum_{\ell=0}^{k} \frac{1}{\ell!}(D^\ell f)(a)(\underbrace{x-a, \cdots, x-a}_{\ell \text{ copies } x-a}).$$

Corollary 5.70. *Let $U \subseteq \mathbb{R}^n$ be open, $f : U \to \mathbb{R}$ be $(k+1)$-times differentiable, and define the remainder*

$$R_k(a,h) = f(a+h) - \sum_{\ell=0}^{k} \frac{1}{\ell!}(D^\ell f)(a)(h,\cdots,h).$$

Then $\lim\limits_{h\to 0} \dfrac{R_k(a,h)}{\|h\|_{\mathbb{R}^n}^k} = 0$, *or in notation,* $R_k(a,h) = o(\|h\|_{\mathbb{R}^n}^k)$ *as* $h \to 0$.

Remark 5.71. An n-dimensional multi-index is a vector $\alpha = (\alpha_1, \cdots, \alpha_n)$ of non-negative integers (that is, $\alpha_j \in \mathbb{N} \cup \{0\}$ for all $1 \leqslant j \leqslant n$). Given an n-dimensional multi-index $\alpha = (\alpha_1, \cdots, \alpha_n)$, $|\alpha|$ and $\alpha!$ are numbers defined by

$$|\alpha| = \sum_{k=1}^{n} \alpha_k \quad \text{and} \quad \alpha! = \prod_{k=1}^{n} \alpha_k!,$$

and the differential operator D_x^α is defined by

$$D_x^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

We also use D^α to denote D_x^α when the variable of differentiation is clear. For a vector $h = (h_1, \cdots, h_n) \in \mathbb{R}^n$ and an n-dimensional multi-index α, we use h^α to denote the number $h_1^{\alpha_1} h_2^{\alpha_2} \cdots h_n^{\alpha_n}$.

Suppose that $f : U \to \mathbb{R}$ is $(k+1)$-times differentiable. Then $D^\ell f$ is continuous on U for $1 \leqslant \ell \leqslant k$; that is, f is of class $\mathscr{C}^k$; thus Theorem 5.65 implies that all the mixed partial derivatives $\dfrac{\partial^\ell f}{\partial x_{j_\ell} \partial x_{j_{\ell+1}} \cdots \partial x_{j_1}}$ are continuous on U. Therefore, the Clairaut Theorem shows that

$$(D^\ell f)(x)(h,\cdots,h) = \sum_{|\alpha|=\ell} \frac{|\alpha|!}{\alpha!}(D^\alpha f)(x) h^\alpha \qquad \forall\, x \in U, h \in \mathbb{R}^n,$$

and the Taylor Theorem further implies that

$$f(x) = \sum_{\ell=0}^{k} \sum_{|\alpha|=k} \frac{1}{\alpha!}(D^\alpha f)(a)(x-a)^\alpha + \sum_{|\alpha|=m+1} \frac{1}{\alpha!}(D^\alpha f)(c)(x-a)^\alpha.$$

Example 5.72. Let $f(x,y) = e^x \cos y$. Compute the fourth degree Taylor polynomial for f at $(0,0)$.

Solution: We compute the zeroth, the first, the second, the third and the fourth mixed derivatives of f at $(0,0)$ as follows:

$$\begin{aligned}
f(0,0) &= 1, \qquad f_x(0,0) = 1, \qquad f_y(0,0) = 0, \\
f_{xx}(0,0) &= 1, \qquad f_{xy}(0,0) = f_{yx}(0,0) = 0, \qquad f_{yy}(0,0) = -1, \\
f_{xxx}(0,0) &= 1, \qquad f_{xxy}(0,0) = f_{xyx}(0,0) = f_{yxx}(0,0) = 0, \\
f_{yyy}(0,0) &= 0, \qquad f_{yyx}(0,0) = f_{yxy}(0,0) = f_{xyy}(0,0) = -1,
\end{aligned}$$

and

$$\begin{aligned}
& f_{xxxx}(0,0) = 1, \qquad f_{yyyy}(0,0) = 1, \\
& f_{xxxy}(0,0) = f_{xxyx}(0,0) = f_{xyxx}(0,0) = f_{yxxx}(0,0) = 0, \\
& f_{xyyy}(0,0) = f_{yxyy}(0,0) = f_{yyxy}(0,0) = f_{yyyx}(0,0) = 0, \\
& f_{xxyy}(0,0) = f_{xyxy}(0,0) = f_{xyyx}(0,0) = f_{yxxy}(0,0) = f_{yxyx}(0,0) = f_{yyxx}(0,0) = -1.
\end{aligned}$$

Then the fourth order Taylor polynomial for f at $(0,0)$ is

$$\begin{aligned}
& f(0,0) + f_x(0,0)x + f_y(0,0)y + \frac{1}{2}\Big[f_{xx}(0,0)x^2 + 2f_{xy}(0,0)xy + f_{yy}(0,0)y^2\Big] \\
& \quad + \frac{1}{6}\Big[f_{xxx}(0,0)x^3 + 3f_{xxy}(0,0)x^2y + 3f_{xyy}(0,0)xy^2 + f_{yyy}(0,0)y^3\Big] \\
& \quad + \frac{1}{24}\Big[f_{xxxx}(0,0)x^4 + 4f_{xxxy}(0,0)x^3 + 6f_{xxyy}(0,0)x^2y^2 \\
& \qquad\quad + 4f_{xyyy}(0,0)xy^3 + f_{yyyy}(0,0)y^4\Big] \\
& = 1 + x + \frac{1}{2}\big(x^2 - y^2\big) + \frac{1}{6}\big(x^3 - 3xy^2\big) + \frac{1}{24}\big(x^4 - 6x^2y^2 + y^4\big).
\end{aligned}$$

Observing that using the Taylor expansions

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \cdots \quad \text{and} \quad \cos y = 1 - \frac{1}{2}y^2 + \frac{1}{24}y^4 + \cdots,$$

we can "formally" compute $e^x \cos y$ by multiplying the two "polynomials" above and obtain that

$$e^x \cos y \text{ “=” } 1 + x + \frac{1}{2}\big(x^2 - y^2\big) + \big(\frac{1}{6}x^3 - \frac{1}{2}xy^2\big) + \big(\frac{1}{24}x^4 - \frac{1}{4}x^2y^2 + \frac{1}{24}y^2\big) + \text{h.o.t.};$$

where h.o.t. stands for the higher order terms which are terms with fifth or higher degree.

Theorem 5.73. *Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$ be of class $\mathscr{C}^k$ and $(D^\ell f)(a) = 0$ for $\ell = 1, \cdots, k-1$. If $(D^k f)(a)(u, u, \cdots, u) > 0$ for all non-zero vectors $u \in \mathbb{R}^n$, then f has a local minimum at a; that is, there exists $\delta > 0$ such that*

$$f(x) \geqslant f(a) \qquad \forall\, x \in B(a, \delta).$$

Proof. Let $a \in U$. Since U is open, there exists $r > 0$ such that $B(a,r) \subseteq U$. Note that $g : B(a,r) \times \mathbb{R}^n \to \mathbb{R}$ defined by $g(x,u) = (D^k f)(x)(u,\cdots,u)$ is continuous since

$$\begin{aligned}
\big|g(x,u) - g(y,v)\big| &= \big|(D^k f)(x)(u,\cdots,u) - (D^k f)(y)(v,\cdots,v)\big| \\
&\leqslant \big|(D^k f)(x)(u,\cdots,u) - (D^k f)(x)(v,\cdots,v)\big| + \Big|\big[(D^k f)(x) - (D^k f)(y)\big](v,\cdots,v)\Big| \\
&\leqslant \big|(D^k f)(x)(u,\cdots,u) - (D^k f)(x)(v,\cdots,v)\big| + \big|(D^k f)(x) - (D^k f)(y)\big\|\|v\|_{\mathbb{R}^n}^k \\
&\leqslant \big|(D^k f)(x)(u-v,u,\cdots,u)\big| + \big|(D^k f)(x)(v,u,\cdots,u) - (D^k f)(x)(v,\cdots,v)\big| \\
&\quad + \big|(D^k f)(x) - (D^k f)(y)\big\|\|v\|_{\mathbb{R}^n}^k \\
&\leqslant \big\|(D^k f)(x)\big\|\|u-v\|_{\mathbb{R}^n}\|u\|_{\mathbb{R}^n}^{k-1} + \big|(D^k f)(x)(v,u-v,u\cdots,u) - (D^k f)(x)(v,\cdots,v)\big| \\
&\quad + \big|(D^k f)(x)(v,v,u\cdots,u) - (D^k f)(x)(v,\cdots,v)\big| + \big|(D^k f)(x) - (D^k f)(y)\big\|\|v\|_{\mathbb{R}^n}^k \\
&\leqslant \cdots\cdots \\
&\leqslant \big\|(D^k f)(x)\big\|\|u-v\|_{\mathbb{R}^n}\big(\|u\|_{\mathbb{R}^n}^{k-1} + \|u\|_{\mathbb{R}^n}^{k-2}\|v\|_{\mathbb{R}^n} + \cdots + \|u\|_{\mathbb{R}^n}\|v\|_{\mathbb{R}^n}^{k-2} + \|v\|_{\mathbb{R}^n}^{k-1}\big) \\
&\quad + \big\|(D^k f)(x) - (D^k f)(y)\big\|\|v\|_{\mathbb{R}^n}^k
\end{aligned}$$

so that

$$\begin{aligned}
&\big|g(x,u) - g(y,v)\big| \\
&\qquad \leqslant \big\|(D^k f)(x)\big\|\big(\|u\|_{\mathbb{R}^n} + \|v\|_{\mathbb{R}^n}\big)^{k-1}\|u-v\|_{\mathbb{R}^n} + \big\|(D^k f)(x) - (D^k f)(y)\big\|\|v\|_{\mathbb{R}^n}^k
\end{aligned} \tag{5.7.3}$$

and the right-hand side approaches zero as $x \to y$ and $u \to v$. In particular, by the compactness of $\mathbb{S}^{n-1} \equiv \big\{x \in \mathbb{R}^n \,\big|\, \|x\| = 1\big\}$($= B[0,1]\backslash B(0,1)$ which is closed and bounded), $g(a,\cdot)$ attains its minimum at some point $w \in \mathbb{S}^{n-1}$; that is,

$$g(a,u) \geqslant g(a,w) \qquad \forall\, u \in \mathbb{S}^{n-1}.$$

Let $\lambda = g(a,w) = (D^k f)(a)(w,\cdots,w) > 0$. Since f is of class $\mathscr{C}^k$, there exists $0 < \delta < r$ such that

$$\|(D^k f)(x) - (D^k f)(a)\| < \frac{\lambda}{2} \quad \text{whenever} \quad x \in B(a,\delta).$$

Let $x \in B(a,\delta)\backslash\{a\}$ be given. By Taylor's Theorem there exists $c \in \overline{xa}$ (so that $c \in B(a,\delta)$) such that

$$f(x) = f(a) + \sum_{\ell=1}^{k-1}\frac{1}{\ell!}(D^\ell f)(a)(\overbrace{x-a,\cdots,x-a}^{\ell\text{ copies of }x-a}) + \frac{1}{k!}(D^k f)(c)(\overbrace{x-a,\cdots,x-a}^{k\text{ copies of }x-a}).$$

Since $(D^\ell f)(a)(u,u,\cdots,u) = 0$ for $1 \leqslant j \leqslant k-1$, we conclude that

$$f(x) = f(a) + \frac{1}{k!}(D^k f)(c)(x-a,x-a,\cdots,x-a) = f(a) + \frac{1}{k!}g(c,x-a).$$

Note that (5.7.3) implies that

$$\Big|g\big(c, \frac{x-a}{\|x-a\|}\big) - g\big(a, \frac{x-a}{\|x-a\|}\big)\Big| \leqslant \|(D^k f)(c) - (D^k f)(a)\| < \frac{\lambda}{2};$$

thus

$$g\big(c, \frac{x-a}{\|x-a\|}\big) > g\big(a, \frac{x-a}{\|x-a\|}\big) - \frac{\lambda}{2} \geqslant \frac{\lambda}{2}.$$

By the fact that $g(c, x-a) = g\big(c, \frac{x-a}{\|x-a\|}\big)\|x-a\|^k$, we conclude that

$$f(x) > f(a) + \frac{\lambda}{2k!}\|x-a\|^k \qquad \forall\, x \in B(a,\delta)\backslash\{a\};$$

thus $f(x) \geqslant f(a)$ for all $x \in B(a,\delta)$. □

Corollary 5.74. *Let $U \subseteq \mathbb{R}^n$ be open, $a \in U$, and $f : U \to \mathbb{R}$ be of class $\mathscr{C}^2$. If $\dfrac{\partial f}{\partial x_\ell}(a) = 0$ for all $1 \leqslant \ell \leqslant n$ and the Hessian matrix of f at a is positive (cf. negative) definitive, then f has a local minimum (cf. maximum) at a.*

Definition 5.75. Let $U \subseteq \mathbb{R}^n$ be open. A function $f : U \to \mathbb{R}$ is said to be ***real analytic*** at $a \in U$ if $f(x) = \sum\limits_{k=0}^{\infty} \frac{1}{k!}(D^k f)(a)(x-a, \cdots, x-a)$ in a neighborhood of a.

Example 5.76. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} \exp\big(-\dfrac{1}{|x|^2}\big) & \text{if } x > 0, \\ 0 & \text{if } x < 0. \end{cases}$$

Then f is of class $\mathscr{C}^\infty$, and $f^{(k)}(0) = 0$ for all $k \in \mathbb{N}$. Therefore, f is not real analytic at 0.

5.8 Exercises

§ 5.1 Bounded Linear Maps

Problem 5.1. Let $\{T_k\}_{k=1}^\infty \subseteq \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$ be a sequence of bounded linear maps from $\mathbb{R}^n \to \mathbb{R}^m$. Prove that the following three statements are equivalent:

1. there exists a function $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $\{T_k \boldsymbol{x}\}_{k=1}^\infty$ converges to $T\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$;

2. $\lim\limits_{k,\ell\to\infty} \|T_k - T_\ell\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} = 0;$

3. there exists a function $T : \mathbb{R}^n \to \mathbb{R}^m$ such that for every compact $K \subseteq \mathbb{R}^n$ and $\varepsilon > 0$ there exists $N > 0$ such that

$$\|T_k\boldsymbol{x} - T\boldsymbol{x}\|_{\mathbb{R}^m} < \varepsilon \quad \text{whenever} \quad \boldsymbol{x} \in K \ \text{ and } \ k \geqslant N\,.$$

Problem 5.2. Recall that $\mathcal{M}_{m\times n}$ is the collection of all $m \times n$ real matrices. For a given $A \in \mathcal{M}_{m\times n}$, define a function $f : \mathcal{M}_{n\times m} \to \mathbb{R}$ by

$$f(M) = \operatorname{tr}(AM)\,,$$

where tr is the trace operator which maps a square matrix to the sum of its diagonal entries. Show that $f \in \mathscr{B}(\mathcal{M}_{n\times m}, \mathbb{R})$.
Hint: You may need the conclusion in Example 4.29.

Problem 5.3. Let $\mathscr{P}([0,1))$ be the collection of all polynomials defined on $[0,1]$, and $\|\cdot\|_\infty$ be the max-norm defined by $\|p\|_\infty = \max\limits_{x\in[0,1]} |p(x)|$.

1. Show that the differential operator $\dfrac{d}{dx} : \mathscr{P}([0,1]) \to \mathscr{P}([0,1])$ is linear.

2. Show that $\dfrac{d}{dx} : \big(\mathscr{P}([0,1]), \|\cdot\|_\infty\big) \to \big(\mathscr{P}([0,1]), \|\cdot\|_\infty\big)$ is unbounded; that is, show that
$$\sup_{\|p\|_\infty=1} \|p'\|_\infty = \infty\,.$$

Problem 5.4. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, and $T \in \mathscr{B}(X,Y)$. Show that for all $\boldsymbol{x} \in X$ and $r > 0$,

$$\sup_{\boldsymbol{x}'\in B(\boldsymbol{x},r)} \|T\boldsymbol{x}'\|_Y \geqslant r\|T\|_{\mathscr{B}(X,Y)}\,.$$

Hint: Prove and make use of the inequality $\max\big\{\|T(\boldsymbol{x}+\boldsymbol{\xi})\|_Y, \|T(\boldsymbol{x}-\boldsymbol{\xi})\|_Y\big\} \geqslant \|T\boldsymbol{\xi}\|_Y$ for all $\boldsymbol{\xi} \in Y$.

Problem 5.5. Let $(X, \|\cdot\|_X)$ be a Banach space, $(Y, \|\cdot\|_Y)$ be a normed space, and $\mathscr{F} \subseteq \mathscr{B}(X,Y)$ be a family of bounded linear maps from X to Y. Show that if $\sup\limits_{T\in\mathscr{F}} \|T\boldsymbol{x}\|_Y < \infty$ for all $x \in X$, then

$$\sup_{T\in\mathscr{F}} \|T\|_{\mathscr{B}(X,Y)} < \infty\,.$$

Hint: Suppose the contrary that there exists $\{T_n\}_{n=1}^{\infty} \subseteq \mathscr{F}$ such that $\|T_n\|_{\mathscr{B}(X,Y)} \geqslant 4^n$. Using Problem 5.4 to choose a sequence $\{\boldsymbol{x}_n\}_{n=0}^{\infty}$, where $\boldsymbol{x}_0 = \boldsymbol{0}$, such that

$$\boldsymbol{x}_n \in B(\boldsymbol{x}_{n-1}, 3^{-n}) \qquad \text{and} \qquad \|T_n \boldsymbol{x}_n\|_Y \geqslant \frac{2}{3} \cdot 3^{-n} \|T_n\|_{\mathscr{B}(X,Y)} \,.$$

Show that $\{\boldsymbol{x}_n\}_{n=1}^{\infty}$ converges to some point $\boldsymbol{x} \in X$ but $\{T_n \boldsymbol{x}\}_{n=1}^{\infty}$ is not bounded in Y.
Remark: The conclusion above is called the Uniform Boundedness Principle (or the Banach-Steinhaus Theorem). This is one of the fundamental results in functional analysis.

§ 5.2 Definition of Derivatives

Problem 5.6. Show that if $f : \mathbb{C} \to \mathbb{R}$ is differentiable at z_0, then $(Df)(z_0) = 0$.
Hint: Show that $\mathscr{B}(\mathbb{C}, \mathbb{R}) = \{0\}$.

Problem 5.7. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x,y) = \begin{cases} 0 & \text{if } \; xy = 0\,, \\ 1 & \text{if } \; xy \neq 0\,. \end{cases}$$

Compute $\dfrac{\partial f}{\partial x}(x,y)$ and $\dfrac{\partial f}{\partial y}(x,y)$.

Problem 5.8. Investigate the differentiability of

$$f(x,y) = \begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

Problem 5.9. Investigate the differentiability of

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

Problem 5.10. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

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

Discuss the differentiability of f. Find $(\nabla f)(x,y)$ at points of differentiability.

Problem 5.11. Let $(X, \|\cdot\|_X)$ be a normed space, $A \subseteq X$ and $f : A \to \mathbb{R}$ be a function. f is said to attain a local extremum at $a \in A$ if there exists $r > 0$ such that either $f(x) \leqslant f(a)$ for all $x \in B(a,r) \cap A$ or $f(x) \geqslant f(a)$ for all $x \in B(a,r) \cap A$. Show that if f attains a local extremum at an interior point $a \in \mathring{A}$ and f is differentiable at a, then $(Df)(a)$ is the zero map in $\mathscr{B}(X, \mathbb{R})$.

Problem 5.12. Let $X = \mathcal{M}_{n\times m}$, the collection of all $n \times m$ real matrices, equipped with the Frobenius norm $\|\cdot\|_F$ introduced in Exercise Problem 2.9, and $f : X \to \mathbb{R}$ be defined by $f(A) = \|A\|_F^2$. Show that f is differentiable on X and find $(Df)(A)$ for $A \in X$.

Problem 5.13. Let $\|\cdot\|_F$ denote the Frobenius norm of matrices given in Exercise Problem 2.9. For an $m \times n$ matrix $A = [a_{ij}]$, we look for an $m \times k$ matrix $C = [c_{ij}]$ and an $k \times n$ matrix $R = [r_{ij}]$, where $1 \leqslant k \leqslant \min\{m, n\}$, such that $\|A - CR\|_F^2$ is minimized. This is to minimize the function

$$f(C,R) = \|A - CR\|_F^2 = \mathrm{tr}((A - CR)(A - CR)^{\mathrm{T}}) = \sum_{i=1}^{n}\sum_{j=1}^{m}\Big(a_{ij} - \sum_{\ell=1}^{k} c_{i\ell} r_{\ell j}\Big)^2 .$$

Show that if $C \in \mathbb{R}^{m\times k}$ and $R \in \mathbb{R}^{k\times n}$ minimize f, then C, R satisfy

$$(A - CR)R^{\mathrm{T}} = 0 \qquad \text{and} \qquad C^{\mathrm{T}}(A - CR) = 0 .$$

Problem 5.14. Let $u, v \in \mathbb{R}^n$ be given vectors satisfying $u \cdot v \neq 0$, and $B \in \mathcal{M}_{n\times n}(\mathbb{R})$ be a symmetric matrix. Define $f : \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathbb{R}$ be defined by

$$f(X) = \|X - B\|_F^2 = \mathrm{tr}((X - B)^{\mathrm{T}}(X - B)) .$$

1. Find the derivative of f.

2. Consider finding the minimum of f over the set

$$S \equiv \big\{X \in \mathcal{M}_{n\times n}(\mathbb{R}) \,\big|\, X = X^{\mathrm{T}} \text{ and } Xu = v\big\} .$$

   Suppose that A is the minimizer of f over S; that is,

$$f(A) \leqslant f(X) \qquad \forall\, X \in S .$$

   Show that

3.

Proof. 1. Let $H \in \mathcal{M}_{n\times n}(\mathbb{R})$ be given. Then

$$\begin{aligned} f(A+H) - f(A) &= \operatorname{tr}((A+H-B)^{\mathrm{T}}(A+H-B)) - \operatorname{tr}((A-B)^{\mathrm{T}}(A-B)) \\ &= \operatorname{tr}(H^{\mathrm{T}}(A-B) + (A-B)^{\mathrm{T}}H) + \operatorname{tr}(H^{\mathrm{T}}H)\,. \end{aligned}$$

Define $L(H) = \operatorname{tr}(H^{\mathrm{T}}(A-B) + (A-B)^{\mathrm{T}}H)$. Then $L : \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathbb{R})$ is linear. Moreover, for $H \neq 0$ we have

$$\frac{\big|f(A+H) - f(A) - L(H)\big|}{\|H\|_F} = \frac{\big|\operatorname{tr}(H^{\mathrm{T}}H)\big|}{\|H\|_F} = \frac{\|H\|_F^2}{\|H\|_F} = \|H\|_F\,;$$

thus

$$\lim_{H\to 0} \frac{\big|f(A+H) - f(A) - L(H)\big|}{\|H\|_F} = 0\,.$$

This shows that f is differentiable at A with derivative

$$Df(A)(H) = L(H) = \operatorname{tr}(H^{\mathrm{T}}(A-B) + (A-B)^{\mathrm{T}}H)\,.$$

Since $\operatorname{tr}(M) = \operatorname{tr}(M^{\mathrm{T}})$, we find that $\operatorname{tr}(H^{\mathrm{T}}(A-B) + (A-B)^{\mathrm{T}}H) = 2\operatorname{tr}((A-B)^{\mathrm{T}}H)$; thus $Df(A)(H) = 2\operatorname{tr}((A-B)^{\mathrm{T}}H)$.

2. Let A be the minimizer of f over S. Then

$$\forall\, B \in S, B - A \in \mathcal{N} \equiv \big\{X \in \mathcal{M}_{n\times n}(\mathbb{R}) \,\big|\, X = X^{\mathrm{T}} \text{ and } Xu = 0\big\}.$$

Therefore, any element in S takes the form $A + H$ for some $H \in \mathcal{N}$.

Since A is the minimizer of f over S, we have

$$f(A+H) - f(A) \geqslant 0 \qquad \forall\, H \in \mathcal{N}\,.$$

The computation in part 1 shows that

$$2\operatorname{tr}((A-B)^{\mathrm{T}}H) + \|H\|_F^2 \geqslant 0 \qquad \forall\, H \in \mathcal{N}\,;$$

thus by the fact that $\|H\|_F^2$ is quadratic in $\|H\|$ we must have

$$\operatorname{tr}((A-B)^{\mathrm{T}}H) = 0 \qquad \forall\, H \in \mathcal{N}\,.$$

Let O be a rotation/orthogonal matrix satisfying $\mathrm{O}u = \mathrm{e}_n = (0, \cdots, 0, 1)^{\mathrm{T}}$. Note that

$$H \in \mathcal{N} \quad \Leftrightarrow \quad (\mathrm{O}H\mathrm{O}^{\mathrm{T}}) \text{ is symmetric and } \mathrm{O}H\mathrm{O}^{\mathrm{T}}\mathrm{e}_n = 0\,,$$

and conditions $\mathrm{O}H\mathrm{O}^{\mathrm{T}}$ is symmetric and $\mathrm{O}H\mathrm{O}^{\mathrm{T}}\mathrm{e}_n = 0$ imply that

$$\mathrm{O}H\mathrm{O}^{\mathrm{T}} = \begin{bmatrix} C & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & 0 \end{bmatrix} \quad \text{for some } (n-1) \times (n-1) \text{ symmetric matrix } C.$$

Therefore, the minimizer A satisfies that

$$\operatorname{tr}(\mathrm{O}(A-B)^{\mathrm{T}}\mathrm{O}^{\mathrm{T}}M) = 0 \qquad \forall\, M = \begin{bmatrix} C & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & 0 \end{bmatrix},\, C \in \mathcal{M}_{(n-1)\times(n-1)}(\mathbb{R}) \text{ symmetric}$$

Since $A-B$ is symmetric, the identity above shows that the upper left $(n-1)\times(n-1)$ block of $\mathrm{O}(A-B)\mathrm{O}^{\mathrm{T}}$ is the zero matrix; thus

$$\mathrm{O}(A-B)\mathrm{O}^{\mathrm{T}} = \begin{bmatrix} 0 & \cdots & 0 & a_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n-1} \\ a_1 & \cdots & a_{n-1} & a_n \end{bmatrix}$$

which further shows that

$$A - B = \mathrm{O}^{\mathrm{T}} \begin{bmatrix} 0 & \cdots & 0 & a_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n-1} \\ a_1 & \cdots & a_{n-1} & a_n \end{bmatrix} \mathrm{O}$$

3.

□

Problem 5.15. Let $X = \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$ equipped with norm $\|\cdot\|$, and $f : \mathrm{GL}(n) \to \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$ be defined by $f(L) = L^{-2} \equiv L^{-1} \circ L^{-1}$. Show that f is differentiable on $\mathrm{GL}(n)$ and find $(Df)(L)$ for $L \in \mathrm{GL}(n)$.

Problem 5.16. Let $X = \mathscr{C}([-,1,1];\mathbb{R})$ and $\|\cdot\|_X$ be defined by $\|f\|_X = \max\limits_{x\in[-1,1]} \big|f(x)\big|$, and $(Y, \|\cdot\|_Y) = (\mathbb{R}, |\cdot|)$. Consider the map $\delta : X \to \mathbb{R}$ be defined by $\delta(f) = f(0)$. Show that δ is differentiable on X. Find $(D\delta)(f)$ (for $f \in \mathscr{C}([-1,1];\mathbb{R})$).

Problem 5.17. Let $X = \mathscr{C}([a,b];\mathbb{R})$ and $\|\cdot\|_2$ be the norm induced by the inner product $\langle f, g\rangle = \int_a^b f(x)g(x)\,dx$. Define $I : X \to X$ by

$$I(f)(x) = \int_a^x f(t)^2\,dt \qquad \forall\, x \in [a,b]\,.$$

Show that I is differentiable on X, and find $(DI)(f)$.

Problem 5.18. Let $X = \mathscr{D}([a,b];\mathbb{R})$, the collection of all piecewise continuously differentiable real-valued function, and $\|\cdot\|_2$ be the norm (on X) induced by the inner product

$$\langle f, g\rangle = \int_a^b \big[f(x)g(x) + f'(x)g'(x)\big]\,dx\,.$$

For $g \in \mathscr{C}([a,b];\mathbb{R})$, define $I : X \to \mathbb{R}$ by

$$I(f) = \int_a^b \big[g(t) - f'(t)\big]^2\,dt\,.$$

Show that I is differentiable on X, and find $(DI)(f)$.

Problem 5.19. Let $r > 0$ and $\alpha > 1$. Suppose that $f : B(0,r) \to \mathbb{R}$ satisfies $|f(x)| \leqslant \|x\|^\alpha$ for all $x \in B(0,r)$. Show that f is differentiable at 0. What happens if $\alpha = 1$?

Problem 5.20. Suppose that $f, g : \mathbb{R} \to \mathbb{R}^m$ are differentiable at a and there is a $\delta > 0$ such that $g(x) \neq 0$ for all $0 < |x - a| < \delta$. If $f(a) = g(a) = 0$ and $(Dg)(a) \neq 0$, show that

$$\lim_{x\to a} \frac{\|f(x)\|}{\|g(x)\|} = \frac{\|(Df)(a)\|}{\|(Dg)(a)\|}\,.$$

In the following, we discuss the concept of sub-differentials of convex functions. Note that convex functions may not be differentiable everywhere, and sub-differentials extends the idea of derivatives.

Definition 5.77. Let $C \subseteq \mathbb{R}^n$ be a convex set, $f : C \to \mathbb{R}$ be a convex function, and $\boldsymbol{x} \in C$. A vector $\boldsymbol{v} \in \mathbb{R}^n$ is called a sub-gradient of f at $\boldsymbol{x}$ if

$$f(\boldsymbol{y}) \geqslant f(\boldsymbol{x}) + \boldsymbol{v}^\top(\boldsymbol{y} - \boldsymbol{x}) \qquad \forall\, \boldsymbol{y} \in C\,. \tag{5.8.1}$$

and the sub-differential of f at $\boldsymbol{x}$, denoted by $\partial f(\boldsymbol{x})$, is the collection of all sub-gradients of f at $\boldsymbol{x}$; that is,

$$\partial f(\boldsymbol{x}) = \big\{\boldsymbol{v} \in \mathbb{R}^n \,\big|\, f(\boldsymbol{y}) \geqslant f(\boldsymbol{x}) + \boldsymbol{v}^\top(\boldsymbol{y} - \boldsymbol{x}) \text{ for all } \boldsymbol{y} \in C\big\}.$$

We note that if in addition f is indeed differentiable, then the sub-differential of f at $\boldsymbol{x}$ contains only the vector $(Df)(\boldsymbol{x})$.

Problem 5.21. Let $C \subseteq \mathbb{R}^n$ be a convex set, $f : C \to \mathbb{R}$ be a convex function, and $\boldsymbol{x} \in \text{int}(C)$.

1. Show that the sub-differential $\partial f(\boldsymbol{x})$ is non-empty, convex and compact.

2. Show that f attains a local minimum at $\boldsymbol{x}$ if and only if $\mathbf{0} \in \partial f(\boldsymbol{x})$.

Problem 5.22. Let $C \subseteq \mathbb{R}^n$ be a convex set, and $f, g : C \to \mathbb{R}$ be convex functions. If in addition that g is differentiable (in the interior of C), then for $\boldsymbol{x} \in \text{int}(C)$,

$$\boldsymbol{v} \in \partial(f+g)(\boldsymbol{x}) \qquad \text{if and only if} \qquad \boldsymbol{v} - (Dg)(\boldsymbol{x}) \in \partial f(\boldsymbol{x}).$$

§ 5.3 Continuity of Differentiable Maps

Problem 5.23. Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$. Suppose that the partial derivatives $\dfrac{\partial f}{\partial x_1}, \cdots, \dfrac{\partial f}{\partial x_n}$ are bounded on U; that is, there exists a real number $M > 0$ such that

$$\Big|\frac{\partial f}{\partial x_j}(x)\Big| \leqslant M \qquad \forall\, x \in U \text{ and } j = 1, \cdots, n\,.$$

Show that f is continuous on U.
Hint: Mimic the proof of Theorem 5.40.

§ 5.4 Conditions for Differentiability

Problem 5.24. Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$. Show that f is differentiable at $a \in U$ if and only if there exists a vector-valued function $\varepsilon : U \to \mathbb{R}^n$ such that

$$f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j) = \varepsilon(x) \cdot (x - a)$$

and $\varepsilon(x) \to 0$ as $x \to a$.

Problem 5.25. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} \dfrac{x^3 y}{x^4 + y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

and $u \in \mathbb{R}^2$ be a unit vector. Show that the directional derivative of f at the origin exists in all direction, and

$$(D_u f)(0,0) = \Big(\frac{\partial f}{\partial x}(0,0), \frac{\partial f}{\partial y}(0,0)\Big) \cdot u\,.$$

Is f differentiable at $(0,0)$?

§ 5.5 The Product Rule and the Chain Rule

Problem 5.26. Verify the chain rule for

$$u(x,y,z) = xe^y, \quad v(x,y,z) = yz\sin x$$

and

$$f(u,v) = u^2 + v\sin u$$

with $h(x,y,z) = f\big(u(x,y,z), v(x,y,z)\big)$.

Problem 5.27. Let (r,θ,φ) be the spherical coordinate of $\mathbb{R}^3$ so that

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

1. Find the Jacobian matrices of the map $(x,y,z) \mapsto (r,\theta,\varphi)$ and the map $(r,\theta,\varphi) \mapsto (x,y,z)$.

2. Suppose that $f(x,y,z)$ is a differential function in $\mathbb{R}^3$. Express $|\nabla f|^2$ in terms of the spherical coordinate.

Problem 5.28 (The inverse statement of the chain rule)**.** Let $f : (a,b) \to \mathbb{R}$ be continuous and $g : (c,d) \to \mathbb{R}$ be differentiable at $y_0 = f(x_0) \in (c,d)$. Show that if $(g \circ f)$ is differentiable at x_0 and $g'(y_0) \neq 0$, then f is differentiable at x_0.

Problem 5.29. Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}^m$ with $f = (f_1, \cdots, f_m)$.

1. Suppose that f is differentiable on U and the line segment joining x and y lies in U. Then there exist points $c_1, \cdots, c_m$ on that segment such that

$$f_i(y) - f_i(x) = (Df_i)(c_i)(y-x) \qquad \forall\, i = 1, \cdots, m.$$

2. Suppose in addition that U is convex (the convexity of sets is defined in Problem 4.16). Show that for each $x, y \in U$ and vector $v \in \mathbb{R}^m$, there exists c on the line segment joining x and y such that

$$v \cdot \big[f(x) - f(y)\big] = v \cdot (Df)(c)(x-y)\,.$$

In particular, show that if $\sup\limits_{x\in U} \|(Df)(x)\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} \leqslant M$, then

$$\|f(x) - f(y)\|_{\mathbb{R}^m} \leqslant M\|x-y\|_{\mathbb{R}^n} \qquad \forall\, x, y \in U\,.$$

Problem 5.30. Let $U \subseteq \mathbb{R}^n$ be open and connected, and $f : U \to \mathbb{R}$ be a function such that $\dfrac{\partial f}{\partial x_j}(x) = 0$ for all $x \in U$. Show that f is constant in U.

Problem 5.31. Let $U \subseteq \mathbb{R}^n$ be open, and for each $1 \leqslant i, j \leqslant n$, $a_{ij} : U \to \mathbb{R}$ be differentiable functions. Define $A = [a_{ij}]$ and $J = \det(A)$. Show that

$$\frac{\partial J}{\partial x_k} = \operatorname{tr}\Big(\operatorname{Adj}(A)\frac{\partial A}{\partial x_k}\Big) \qquad \forall\, 1 \leqslant k \leqslant n\,,$$

where for a square matrix $M = [m_{ij}]$, $\operatorname{tr}(M)$ denotes the trace of M, $\operatorname{Adj}(M)$ denotes the adjoint matrix of M, and $\dfrac{\partial M}{\partial x_k}$ denotes the matrix whose (i,j)-th entry is given by $\dfrac{\partial m_{ij}}{\partial x_k}$.

Hint: Show that

$$\frac{\partial J}{\partial x_k} = \begin{vmatrix} \frac{\partial a_{11}}{\partial x_k} & a_{12} & \cdots & a_{1n} \\ \frac{\partial a_{21}}{\partial x_k} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ \frac{\partial a_{n1}}{\partial x_k} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \frac{\partial a_{12}}{\partial x_k} & a_{13} & \cdots & a_{1n} \\ a_{21} & \frac{\partial a_{22}}{\partial x_k} & a_{23} & \cdots & a_{2n} \\ \vdots & & & & \vdots \\ a_{n1} & \frac{\partial a_{n2}}{\partial x_k} & a_{n3} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & \cdots & a_{(n-1)1} & \frac{\partial a_{1n}}{\partial x_k} \\ a_{21} & \cdots & a_{(n-1)2} & \frac{\partial a_{2n}}{\partial x_k} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & a_{(n-1)n} & \frac{\partial a_{n1}}{\partial x_k} \end{vmatrix}$$

and rewrite this identity in the form which is asked to prove. You can also show the differentiation formula by applying the chain rule to the composite function $F \circ g$ of maps $g : U \to \mathbb{R}^{n^2}$ and $F : \mathbb{R}^{n^2} \to \mathbb{R}$ defined by $g(x) = \big(a_{11}(x), a_{12}(x), \cdots, a_{nn}(x)\big)$ and $F(a_{11}, \cdots, a_{nn}) = \det\big([a_{ij}]\big)$. Check first what $\dfrac{\partial F}{\partial a_{ij}}$ is.

§ 5.6 Higher Derivatives of Functions

Problem 5.32. Let $f(x,y,z) = (x^2+1)\cos(yz)$, and $a = (0, \frac{\pi}{2}, 1)$, $u = (1,0,0)$, $v = (0,1,0)$ and $w = (2,0,1)$.

1. Compute $(Df)(a)(u)$.

2. Compute $(D^2f)(a)(v)(u)$.

3. Compute $(D^3f)(a)(w)(v)(u)$.

Problem 5.33. 1. If $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $g : B \subseteq \mathbb{R}^m \to \mathbb{R}^\ell$ are twice differentiable and $f(A) \subseteq B$, then for $x_0 \in A$, $u, v \in \mathbb{R}^n$, show that

$$\begin{aligned} &D^2(g \circ f)(x_0)(u, v) \\ &\qquad = (D^2g)(f(x_0))\big((Df)(x_0)(u), Df(x_0)(v)\big) + (Dg)(f(x_0))\big((D^2f)(x_0)(u, v)\big) . \end{aligned}$$

2. If $p : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map plus some constant; that is, $p(x) = Lx + c$ for some $L \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$, and $f : A \subseteq \mathbb{R}^m \to \mathbb{R}^s$ is k-times differentiable, prove that

$$D^k(f \circ p)(x_0)(u^{(1)}, \cdots, u^{(k)}) = (D^kf)\big(p(x_0)\big)\big((Dp)(x_0)(u^{(1)}), \cdots, (Dp)(x_0)(u^{(k)}\big) .$$

Problem 5.34. Let $f(x, y)$ be a real-valued function on $\mathbb{R}^2$. Suppose that f is of class $\mathscr{C}^1$ (that is, all first partial derivatives are continuous on $\mathbb{R}^2$) and $\dfrac{\partial^2 f}{\partial x \partial y}$ exists and is continuous. Show that $\dfrac{\partial^2 f}{\partial y \partial x}$ exists and $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$.

Hint: Mimic the proof of Clairaut's Theorem.

Problem 5.35. Let $U \subseteq \mathbb{R}^n$ be open, and $\psi : U \to \mathbb{R}^n$ be a function of class $\mathscr{C}^2$. Suppose that $(D\psi)(x) \in \mathrm{GL}(n)$ for all $x \in \mathbb{R}^n$, and define $J = \det([D\psi])$ and $A = [D\psi]^{-1}$, where $[D\psi]$ is the Jacobian matrix of ψ. Write $[A] = [a_{ij}]$.

1. Show that for each $1 \leqslant i, j, k \leqslant n$, $a_{ij} : \mathbb{R}^n \to \mathbb{R}$ is differentiable, and

$$\frac{\partial a_{ij}}{\partial x_k} = -\sum_{r,s=1}^{n} a_{ir} \frac{\partial^2 \psi_r}{\partial x_k \partial x_s} a_{sj} .$$

2. Show the Piola identity

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i}(Ja_{ij})(x) = 0 \qquad \forall\, 1 \leqslant j \leqslant n \text{ and } x \in U . \tag{5.8.2}$$

§ 5.7 Taylor's Theorem

Problem 5.36. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable, and Df is a constant map in $\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$; that is, $(Df)(x)(u) = (Df)(y)(u)$ for all $x, y \in \mathbb{R}^n$ and $u \in \mathbb{R}^n$. Show that f is a linear term plus a constant and that the linear part of f is the constant value of Df.

Problem 5.37. Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$ be of class $\mathscr{C}^k$ and $(D^\ell f)(a) = 0$ for $\ell = 1, \cdots, k-1$. Show that if $(D^k f)(a)(u, u, \cdots, u) > 0$ for all non-zero vectors $u \in \mathbb{R}^n$, then f has a local minimum at a; that is, there exists $\delta > 0$ such that

$$f(x) \geqslant f(a) \qquad \forall\, x \in B(a, \delta)\,.$$

Problem 5.38. Let $C \subseteq \mathbb{R}^n$ be an open convex set (see Problem 4.16 for the definition of convex sets), and $f : C \to \mathbb{R}$ be a twice differentiable function that attains a local minimum at $x_* \in C$. Suppose that there is a sequence $\{x_k\}_{k=1}^\infty$ in C satisfying $f(x_k) \geqslant f(x_{k+1})$ for all $k \in \mathbb{N}$ and $\liminf\limits_{k\to\infty} \|(\nabla f)(x_k)\| = 0$. Show that if f is strongly convex on C; that is, there exists $m > 0$ such that

$$(D^2 f)(x)(z, z) \geqslant m\|z\|^2 \qquad \forall\, z \in \mathbb{R}^n \text{ and } x \in C\,, \tag{5.8.3}$$

then $\{x_k\}_{k=1}^\infty$ converges to x_*.

§ 5.8 Chapter Review

Problem 5.39 (True or False). Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. Let $U \subseteq \mathbb{R}^n$ be open. Then $f : U \to \mathbb{R}$ is differentiable at $a \in U$ if and only if each directional derivative $(D_u f)(a)$ exists and

   $$(D_u f)(a) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) u_j = \Big(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a)\Big) \cdot u$$

   where $u = (u_1, \cdots, u_n)$ is a unit vector.

2. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be of class $\mathscr{C}^1$. Assume that all second order partial derivatives of f exist, then f is second times differentiable in $\mathbb{R}^2$.

3. Let f be a function defined on $\mathbb{R}^2$, and A be an invertible matrix. Define $y = Ax$ for $x \in \mathbb{R}^n$. Then $f(y)$ is differentiable if and only if $f(Ax)$ is differentiable as a function of x.

4. Let $f : [a, b] \to \mathbb{R}^2$ be continuous and be differentiable on (a, b). If $f(a) = f(b)$, then there exists some $c \in (a, b)$ such that $f'(c) = 0$.

Chapter 6

Integration of Functions

6.1 Integrable Functions

In this chapter, we discuss the integration of (bounded) real-valued functions defined on bounded sets. We first recall the integral of functions of one variable that we learned from Calculus.

Definition 6.1. A finite set $\mathcal{P} = \{x_0, x_1, \cdots, x_n\}$ is called a partition of the closed interval $[a, b]$ if $a = x_0 < x_1 < \cdots < x_n = b$. Such a partition $\mathcal{P}$ is usually denoted by $\{a = x_0 < x_1 < \cdots < x_n = b\}$. The norm of $\mathcal{P}$, denoted by $\|\mathcal{P}\|$, is the number $\max\big\{x_i - x_{i-1} \,\big|\, 1 \leqslant i \leqslant n\big\}$.

Let $f : [a, b] \to \mathbb{R}$ be a function. A Riemann sum of f for the partition $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ of $[a, b]$ is a sum which takes the form

$$\sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}),$$

where $\xi_k \in [x_{k-1}, x_k]$ for each $1 \leqslant k \leqslant n$. f is said to be Riemann integrable on $[a, b]$ if there exists a real number A such that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathcal{P}$ is partition of $[a, b]$ satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sums for the partition $\mathcal{P}$ belongs to the interval $(A - \varepsilon, A + \varepsilon)$. Such a number A (is unique and) is called the Riemann integral of f on $[a, b]$ and is denoted by $\displaystyle\int_{[a,b]} f(x)\,dx$.

To define the partition of a bounded set A in $\mathbb{R}^n$, we start with the simplest case $n = 1$.

Definition 6.2. Let $A \subseteq \mathbb{R}$ be a bounded set. A collection of point $\mathcal{P} = \{x_0, x_1, \cdots, x_N\}$ is called a partition of A if $\mathcal{P}$ is a partition of the closed interval $\big[\inf A, \sup A\big]$. Such a partition $\mathcal{P}$ is usually denoted by $\{[x_k, x_{k+1}] \,|\, 0 \leqslant k \leqslant N-1\}$, and the norm of $\mathcal{P}$ is the maximum of the length of intervals in $\mathcal{P}$; that is,

$$\|\mathcal{P}\| \equiv \max\{x_k - x_{k-1} \,|\, 1 \leqslant k \leqslant N\}.$$

Next we look at how a partition of a bounded set in the plane is defined.

Definition 6.3. Let $A \subseteq \mathbb{R}^2$ be a bounded set. Define

$$\begin{aligned} a_1 &= \inf\{x \in \mathbb{R} \,|\, (x,y) \in A \text{ for some } y \in \mathbb{R}\}, \\ b_1 &= \sup\{x \in \mathbb{R} \,|\, (x,y) \in A \text{ for some } y \in \mathbb{R}\}, \\ a_2 &= \inf\{y \in \mathbb{R} \,|\, (x,y) \in A \text{ for some } x \in \mathbb{R}\}, \\ b_2 &= \sup\{y \in \mathbb{R} \,|\, (x,y) \in A \text{ for some } x \in \mathbb{R}\}. \end{aligned}$$

A collection of rectangles $\mathcal{P}$ is called a ***partition*** of A if there exists a partition $\mathcal{P}_x$ of $[a_1, b_1]$ and a partition P_y of $[a_2, b_2]$, where

$$\mathcal{P}_x = \{a_1 = x_0 < x_1 < \cdots < x_n = b_1\} \quad \text{and} \quad \mathcal{P}_y = \{a_2 = y_0 < y_1 < \cdots < y_m = b_2\},$$

such that

$$\mathcal{P} = \{\Delta_{ij} \,|\, \Delta_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \text{ for } i = 1, 2, \cdots, n \text{ and } j = 1, 2, \cdots, m\}.$$

The norm of $\mathcal{P}$, denoted by $\|\mathcal{P}\|$ and also called the ***mesh size*** of the partition $\mathcal{P}$, is a real number defined by

$$\|\mathcal{P}\| = \max\Big\{\sqrt{(x_i - x_{i-1})^2 + (y_j - y_{j-1})^2} \,\Big|\, i = 1, 2, \cdots, n, j = 1, 2, \cdots, m\Big\}.$$

The number $\sqrt{(x_i - x_{i-1})^2 + (y_j - y_{j-1})^2}$ is often denoted by $\text{diam}(\Delta_{ij})$, and is called the ***diameter*** of Δ_{ij} (thus the norm of $\mathcal{P}$ is the maximum of the diameter of rectangles in $\mathcal{P}$).

In general, the partition of a bounded set $A \subseteq \mathbb{R}^n$ is defined as follows.

Definition 6.4. Let $A \subseteq \mathbb{R}^n$ be a bounded set. Define the numbers $a_1, a_2, \cdots, a_n$ and $b_1, b_2, \cdots, b_n$ by

$$\begin{aligned} a_k &= \inf\{x_k \in \mathbb{R} \,|\, x = (x_1, \cdots, x_n) \in A \text{ for some } x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n \in \mathbb{R}\}, \\ b_k &= \sup\{x_k \in \mathbb{R} \,|\, x = (x_1, \cdots, x_n) \in A \text{ for some } x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n \in \mathbb{R}\}. \end{aligned}$$

A collection of rectangles $\mathcal{P}$ is called a ***partition*** of A if there exists partitions $\mathcal{P}^{(k)}$ of $[a_k, b_k]$, $k = 1, \cdots, n$, $\mathcal{P}^{(k)} = \{a_k = x_0^{(k)} < x_1^{(k)} < \cdots < x_{N_k}^{(k)} = b_k\}$, such that

$$\mathcal{P} = \Big\{\Delta_{i_1 i_2 \cdots i_n} \,\Big|\, \Delta_{i_1 i_2 \cdots i_n} = [x_{i_1-1}^{(1)}, x_{i_1}^{(1)}] \times [x_{i_2-1}^{(2)}, x_{i_2}^{(2)}] \times \cdots \times [x_{i_n-1}^{(n)}, x_{i_n}^{(n+1)}],$$
$$i_k = 1, 2, \cdots, N_k, k = 1, \cdots, n\Big\}.$$

The norm of $\mathcal{P}$, denoted by $\|\mathcal{P}\|$ and also called the ***mesh size*** of the partition $\mathcal{P}$, is a real number defined by

$$\|\mathcal{P}\| = \max\left\{\sqrt{\sum_{k=1}^{n} (x_{i_k}^{(k)} - x_{i_k-1}^{(k)})^2} \,\Big|\, i_k = 1, 2, \cdots, N_k, k = 1, \cdots, n\right\}.$$

The number $\sqrt{\sum\limits_{k=1}^{n} (x_{i_k}^{(k)} - x_{i_{k-1}}^{(k)})^2}$ is often denoted by $\operatorname{diam}(\Delta_{i_1 i_2 \cdots i_n})$, and is called the ***diameter*** of the rectangle $\Delta_{i_1 i_2 \cdots i_n}$. The volume of $\Delta_{i_1 i_2 \cdots i_n}$, denoted by $\nu_n(\Delta_{i_1 i_2 \cdots i_n})$ (or simply $\nu(\Delta_{i_1 i_2 \cdots i_n})$ is n is clear to us), is defined by

$$\nu(\Delta_{i_1 i_2 \cdots i_n}) = \prod_{k=1}^{n} \big(x_{i_k}^{(k)} - x_{i_k-1}^{(k)}\big) = \big(x_{i_1}^{(1)} - x_{i_1-1}^{(1)}\big)\big(x_{i_2}^{(2)} - x_{i_2-1}^{(2)}\big) \cdots \big(x_{i_n}^{(n)} - x_{i_n-1}^{(n)}\big).$$

Next we define the Riemann sum of a function $f : A \to \mathbb{R}$ for a partition $\mathcal{P}$ of A.

Definition 6.5. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a (bounded) function. A ***Riemann sum*** of f for the partition $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$ of A is a sum which takes the form

$$\sum_{k=1}^{N} \overline{f}^A(\xi_k)\nu(\Delta_k),$$

where $\overline{f}^A$ is a function given by

$$\overline{f}^A(x) = \begin{cases} f(x) & x \in A, \\ 0 & x \notin A. \end{cases} \tag{6.1.1}$$

and the set $\Xi = \{\xi_1, \xi_2, \cdots, \xi_N\}$ satisfies that $\xi_k \in \Delta_k$ for all $1 \leqslant k \leqslant N$.

Definition 6.6. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a (bounded) function. The function f is Riemann integrable on A if and only if there exists (a unique) $\mathrm{I} \in \mathbb{R}$ such that for every given $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sums of f for the partition $\mathcal{P}$ belongs to the interval $(\mathrm{I}-\varepsilon, \mathrm{I}+\varepsilon)$.

In other words, f is Riemann integrable on A if and only if there exists $\mathrm{I} \in \mathbb{R}$ such that for every given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\Big|\sum_{k=1}^{N} \overline{f}^{A}(\xi_k)\nu(\Delta_k) - \mathrm{I}\Big| < \varepsilon \tag{6.1.2}$$

whenever $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$ and the set $\Xi = \{\xi_1, \xi_2, \cdots, \xi_N\}$ satisfies that $\xi_k \in \Delta_k$ for all $1 \leqslant k \leqslant N$. The number I is denoted by $(\mathrm{R})\int_A f(x)\,dx$.

The definition of the integrability of functions given above is due to Bernhard Riemann; however, the definition above somehow lacks of flexibility for developing the theory of integration of functions. In the following, we adopt another point of view due to Gaston Darboux to discuss the integration of (bounded) functions $f : A \to \mathbb{R}$ for general bounded set $A \subseteq \mathbb{R}^n$.

Definition 6.7. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a (bounded) function. For a partition $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$, the ***upper sum*** and the ***lower sum*** of f for the partition $\mathcal{P}$, denoted by $U(f,\mathcal{P})$ and $L(f,\mathcal{P})$ respectively, are numbers defined by

$$U(f,\mathcal{P}) = \sum_{k=1}^{N} \sup_{x\in\Delta_k} \overline{f}^{A}(x)\nu(\Delta_k) \qquad \text{and} \qquad L(f,\mathcal{P}) = \sum_{k=1}^{N} \inf_{x\in\Delta_k} \overline{f}^{A}(x)\nu(\Delta_k).$$

The two numbers

$$\overline{\int_A} f(x)\,dx \equiv \inf\big\{U(f,\mathcal{P})\,\big|\,\mathcal{P} \text{ is a partition of } A\big\},$$

and

$$\underline{\int_A} f(x)\,dx \equiv \sup\big\{L(f,\mathcal{P})\,\big|\,\mathcal{P} \text{ is a partition of } A\big\}$$

are called the ***upper integral*** and ***lower integral*** of f on A, respectively. The function f is said to be ***Darboux integrable*** (on A) if $\overline{\int_A} f(x)\,dx$ and $\underline{\int_A} f(x)\,dx$ are identical real numbers, and in this case, we express the upper and lower integral as $(\mathrm{D})\int_A f(x)\,dx$, called the ***Darboux integral*** of f on A.

Using the property of supremum and infimum, we immediately obtain the following

Proposition 6.8. *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f, g : A \to \mathbb{R}$ be functions. If $\mathcal{P}$ is a partition of A, then*

$$L(f, \mathcal{P}) + L(g, \mathcal{P}) \leqslant L(f + g, \mathcal{P}) \leqslant U(f + g, \mathcal{P}) \leqslant U(f, \mathcal{P}) + U(g, \mathcal{P}). \tag{6.1.3}$$

Definition 6.9. A partition $\mathcal{P}'$ of a bounded set $A \subseteq \mathbb{R}^n$ is called a ***refinement*** of another partition $\mathcal{P}$ of A if for any $\Delta' \in \mathcal{P}'$, there is $\Delta \in \mathcal{P}$ such that $\Delta' \subseteq \Delta$. A partition $\mathcal{P}$ of a bounded set $A \subseteq \mathbb{R}^n$ is called the ***common refinement*** of another partitions $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ of A if

1. $\mathcal{P}$ is a refinement of $\mathcal{P}_j$ for all $1 \leqslant j \leqslant k$.

2. If $\mathcal{P}'$ is a refinement of $\mathcal{P}_j$ for all $1 \leqslant j \leqslant k$, then $\mathcal{P}'$ is also a refinement of $\mathcal{P}$.

In other words, $\mathcal{P}$ is a common refinement of $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ if it is the coarsest refinement.

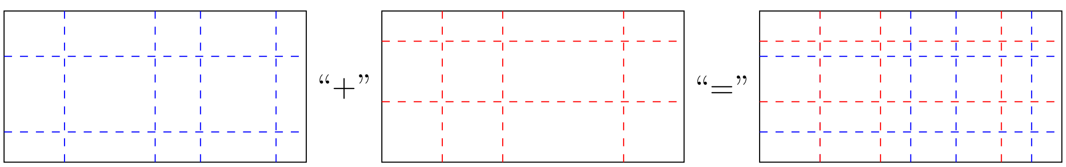


Figure 6.1: The common refinement of two partitions

Qualitatively speaking, $\mathcal{P}$ is a common refinement of $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ if for each $j = 1, \cdots n$, the j-th component c_j of the vertex $(c_1, \cdots, c_n)$ of each rectangle $\Delta \in \mathcal{P}$ belongs to $\mathcal{P}_i^{(j)}$ for some $i = 1, \cdots, k$.

The following proposition should be clear to the readers, and the proof is left as an exercise.

Proposition 6.10. *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a function. If $\mathcal{P}$ and $\mathcal{P}'$ are partitions of A and $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then*

$$L(f, \mathcal{P}) \leqslant L(f, \mathcal{P}') \leqslant U(f, \mathcal{P}') \leqslant U(f, \mathcal{P}).$$

Corollary 6.11. *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a function. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are partitions of A, then $L(f, \mathcal{P}_1) \leqslant U(f, \mathcal{P}_2)$.*

Proof. Let $\mathcal{P}$ be the common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Then Proposition 6.10 implies that

$$L(f, \mathcal{P}_1) \leqslant L(f, \mathcal{P}) \leqslant U(f, \mathcal{P}) \leqslant U(f, \mathcal{P}_2).$$

□

Corollary 6.12. *Let $A \subseteq \mathbb{R}^n$ be a bounded subset, and $f : A \to \mathbb{R}$ be a function. Then*

$$\underline{\int_A} f(x)dx \leqslant \overline{\int_A} f(x)dx .$$

Proof. Note that for each given partition $\mathcal{P}$ of A, the previous corollary implies that $L(f, \mathcal{P})$ is a lower bound for all possible upper sum. Therefore,

$$L(f, \mathcal{P}) \leqslant \overline{\int_A} f(x)dx \qquad \text{for all partitions } \mathcal{P} \text{ of } A$$

which further implies that $\overline{\int_A} f(x)dx$ is an upper bound for all possible lower sum; thus $\underline{\int_A} f(x)dx \leqslant \overline{\int_A} f(x)dx$. □

In the following proposition, we state an equivalent condition for Darboux integrability of bounded functions (on bounded sets).

Proposition 6.13 (Riemann's condition)**.** *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a* (*bounded*) *function. Then f is Darboux integrable on A if and only if*

$$\forall\, \varepsilon > 0, \exists \text{ a partition } \mathcal{P} \text{ of } A \ni U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon .$$

Proof. "⇒" Let $\varepsilon > 0$ be given. By the definition of infimum and supremum, there exist partition $\mathcal{P}_1$ and $\mathcal{P}_2$ of A such that

$$\underline{\int_A} f(x)\, dx - \frac{\varepsilon}{2} < L(f, \mathcal{P}_2) \quad \text{and} \quad \overline{\int_A} f(x)\, dx + \frac{\varepsilon}{2} > U(f, \mathcal{P}_1) .$$

Let $\mathcal{P}$ be a common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Since f is Darboux integrable on A, $\underline{\int_A} f(x)dx = \overline{\int_A} f(x)dx$; thus Proposition 6.10 implies that

$$\begin{aligned} U(f, \mathcal{P}) - L(f, \mathcal{P}) &\leqslant U(f, \mathcal{P}_1) - L(f, \mathcal{P}_2) \\ &< \overline{\int_A} f(x)\, dx + \frac{\varepsilon}{2} - \Big(\underline{\int_A} f(x)\, dx - \frac{\varepsilon}{2} \Big) = \varepsilon . \end{aligned}$$

"⇐" Let $\varepsilon > 0$ be given. By assumption there exists a partition $\mathcal{P}$ of A such that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon$. Then

$$0 \leqslant \overline{\int_A} f(x)\, dx - \underline{\int_A} f(x)\, dx \leqslant U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon .$$

Since $\varepsilon > 0$ is given arbitrary, we must have $\underline{\int_A} f(x)dx = \overline{\int_A} f(x)dx$; thus f is Darboux integrable on A. ▫

The following theorem establishes the equivalence between the Riemann integrals and the Darboux integrals.

Theorem 6.14 (Darboux)**.** *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a (bounded) function. Then f is Riemann integrable on A if and only if f is Darboux integrable on A. In either cases,*

$$(\mathrm{R})\int_A f(x)\,dx = (\mathrm{D})\int_A f(x)\,dx\,.$$

Proof. The boundedness of A guarantees that $A \subseteq \big[-\frac{r}{2}, \frac{r}{2}\big]^n$ for some $r > 0$. Let $R = \big[-\frac{r}{2}, \frac{r}{2}\big]^n$. Then $\nu(R) = r^n$.

"$\Rightarrow$" Suppose that f is Riemann integrable on A with $(\mathrm{R})\int_A f(x)\,dx = \mathrm{I}$. Let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$, then any Riemann of f for $\mathcal{P}$ locates in $\big(\mathrm{I} - \frac{\varepsilon}{4}, \mathrm{I} + \frac{\varepsilon}{4}\big)$.

Let $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ be a partition of A with $\|\mathcal{P}\| < \delta$. For each $1 \leqslant k \leqslant N$, choose $\xi_k, \eta_k \in \Delta_k$ such that

(a) $\sup\limits_{x\in\Delta_k} \overline{f}^A(x) - \dfrac{\varepsilon}{4\nu(R)} < \overline{f}^A(\xi_k) \leqslant \sup\limits_{x\in\Delta_k} \overline{f}^A(x)$;

(b) $\inf\limits_{x\in\Delta_k} \overline{f}^A(x) + \dfrac{\varepsilon}{4\nu(R)} > \overline{f}^A(\eta_k) \geqslant \inf\limits_{x\in\Delta_k} \overline{f}^A(x)$.

Then

$$\begin{aligned} U(f,\mathcal{P}) &= \sum_{k=1}^N \sup_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) < \sum_{k=1}^N \big[\overline{f}^A(\xi_k) + \frac{\varepsilon}{4\nu(R)}\big]\nu(\Delta_k) \\ &= \sum_{k=1}^N \overline{f}^A(\xi_k)\nu(\Delta_k) + \frac{\varepsilon}{4\nu(R)}\sum_{k=1}^N \nu(\Delta_k) < \mathrm{I} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \mathrm{I} + \frac{\varepsilon}{2} \end{aligned}$$

and

$$\begin{aligned} L(f,\mathcal{P}) &= \sum_{k=1}^N \inf_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) > \sum_{k=1}^N \big[\overline{f}^A(\eta_k) - \frac{\varepsilon}{4\nu(R)}\big]\nu(\Delta_k) \\ &= \sum_{k=1}^N \overline{f}^A(\eta_k)\nu(\Delta_k) - \frac{\varepsilon}{4\nu(R)}\sum_{k=1}^N \nu(\Delta_k) > \mathrm{I} - \frac{\varepsilon}{4} - \frac{\varepsilon}{4} = \mathrm{I} - \frac{\varepsilon}{2}\,. \end{aligned}$$

As a consequence, $\mathrm{I}-\frac{\varepsilon}{2} < L(f,\mathcal{P}) \leqslant U(f,\mathcal{P}) < \mathrm{I}+\frac{\varepsilon}{2}$; thus $U(f,\mathcal{P})-L(f,\mathcal{P}) < \varepsilon$ which shows that f is Darboux integrable on A. Moreover, since $\varepsilon > 0$ is given arbitrarily and $L(f,\mathcal{P}) \leqslant (\mathrm{D})\int_A f(x)\,dx \leqslant U(f,\mathcal{P})$, we must have $\mathrm{I} = (\mathrm{D})\int_A f(x)\,dx$.

"$\Leftarrow$" Let $\mathrm{I} = (\mathrm{D})\int_A f(x)dx$, and $\varepsilon > 0$ be given. Since f is Darboux integrable on A, there exists a partition $\mathcal{P}_1$ of A such that $U(f,\mathcal{P}_1) - L(f,\mathcal{P}_1) < \frac{\varepsilon}{2}$. Suppose that $\mathcal{P}_1^{(i)} = \{y_0^{(i)}, y_1^{(i)}, \cdots, y_{m_i}^{(i)}\}$ for $1 \leqslant i \leqslant n$. We define

$$\delta = \frac{\varepsilon}{4r^{n-1}(m_1+m_2+\cdots+m_n+n)\big(\sup \overline{f}^A(R) - \inf \overline{f}^A(R) + 1\big)}\,.$$

Then $\delta > 0$.

Assume that $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$ is a given partition of A with $\|\mathcal{P}\| < \delta$. Let $\mathcal{P}'$ be the common refinement of $\mathcal{P}$ and $\mathcal{P}_1$. Write $\mathcal{P}' = \{\Delta'_1, \Delta'_2, \cdots, \Delta'_{N'}\}$ and $\Delta_k = \Delta_k^{(1)} \times \Delta_k^{(2)} \times \cdots \times \Delta_k^{(n)}$ as well as $\Delta'_k = \Delta_k'^{(1)} \times \Delta_k'^{(2)} \times \cdots \times \Delta_k'^{(n)}$. Define two classes of rectangles in $\mathcal{P}$ and $\mathcal{P}'$ by

$$\begin{aligned} C_1 &= \big\{\Delta \in \mathcal{P} \,\big|\, y_j^{(i)} \notin \Delta^{(i)} \text{ for all } i,j\big\}, \qquad C_2 = \big\{\Delta \in \mathcal{P} \,\big|\, y_j^{(i)} \in \Delta^{(i)} \text{ for some } i,j\big\},\\ D_1 &= \big\{\Delta' \in \mathcal{P}' \,\big|\, y_j^{(i)} \notin \Delta'^{(i)} \text{ for all } i,j\big\}, \quad D_2 = \big\{\Delta \in \mathcal{P}' \,\big|\, y_j^{(i)} \in \Delta'^{(i)} \text{ for some } i,j\big\}. \end{aligned}$$

By the definition of the upper sum,

$$U(f,\mathcal{P}) = \sum_{k=1}^{N} \sup_{x\in\Delta_k} \overline{f}^A(x)\nu(\Delta_k) = \sum_{\Delta\in C_1} \sup_{x\in\Delta} \overline{f}^A(x)\nu(\Delta) + \sum_{\Delta\in C_2} \sup_{x\in\Delta} \overline{f}^A(x)\nu(\Delta)$$

and similarly,

$$U(f,\mathcal{P}') = \sum_{k=1}^{N'} \sup_{x\in\Delta'_k} \overline{f}^A(x)\nu(\Delta'_k) = \sum_{\Delta'\in D_1} \sup_{x\in\Delta'} \overline{f}^A(x)\nu(\Delta') + \sum_{\Delta'\in D_2} \sup_{x\in\Delta'} \overline{f}^A(x)\nu(\Delta')\,.$$

By the fact that $C_1 = D_1$, we must have

$$\sum_{\Delta\in C_1} \sup_{x\in\Delta} \overline{f}^A(x)\nu(\Delta) = \sum_{\Delta'\in D_1} \sup_{x\in\Delta'} \overline{f}^A(x)\nu(\Delta')$$

and

$$\sum_{\Delta\in C_2} \nu(\Delta_k) = \sum_{\Delta'\in D_2} \nu(\Delta')\,.$$

The equalities above further imply that

$$\begin{aligned}
U(f,\mathcal{P})-U(f,\mathcal{P}') &= \sum_{\Delta\in C_2}\sup_{x\in\Delta}\overline{f}^A(x)\nu(\Delta)-\sum_{\Delta'\in D_2}\sup_{x\in\Delta'}\overline{f}^A(x)\nu(\Delta')\\
&\leqslant \big(\sup\overline{f}^A(R)-\inf\overline{f}^A(R)\big)\sum_{\Delta\in C_2}\nu(\Delta)\\
&= \big(\sup\overline{f}^A(R)-\inf\overline{f}^A(R)\big)\sum_{1\leqslant k\leqslant N\text{ with } y_j^{(i)}\in\Delta_k^{(i)}\text{ for some } i,j}\nu(\Delta_k)\\
&= \big(\sup\overline{f}^A(R)-\inf\overline{f}^A(R)\big)\sum_{i=1}^{n}\sum_{j=0}^{m_i}\sum_{1\leqslant k\leqslant N\text{ with } y_j^{(i)}\in\Delta_k^{(i)}}\nu(\Delta_k)\,.
\end{aligned}$$

Moreover, for each fixed i,j,

$$\bigcup_{1\leqslant k\leqslant N\text{ with } y_j^{(i)}\in\Delta_k^{(i)}}\Delta_k\subseteq\Big[-\frac{r}{2},\frac{r}{2}\Big]^{i-1}\times\big[y_j^{(i)}-\delta,y_j^{(i)}+\delta\big]\times\Big[-\frac{r}{2},\frac{r}{2}\Big]^{n-i};$$

thus

$$\sum_{1\leqslant k\leqslant N\text{ with } y_j^{(i)}\in\Delta_k^{(i)}}\nu(\Delta_k)\leqslant 2\delta r^{n-1}\qquad\forall\,1\leqslant i\leqslant n, 1\leqslant j\leqslant m_i\,.$$

Therefore,

$$\begin{aligned}
U(f,\mathcal{P})-U(f,\mathcal{P}') &\leqslant \big(\sup\overline{f}^A(R)-\inf\overline{f}^A(R)\big)\sum_{i=1}^{n}\sum_{j=0}^{m_i}\sum_{1\leqslant k\leqslant N\text{ with } y_j^{(i)}\in\Delta_k^{(i)}}\nu(\Delta_k)\\
&\leqslant \big(\sup\overline{f}^A(R)-\inf\overline{f}^A(R)\big)\sum_{i=1}^{n}\sum_{j=0}^{m_i}2\delta r^{n-1}\\
&\leqslant 2\delta r^{n-1}(m_1+m_2+\cdots+m_n+n)\big(\sup\overline{f}^A(R)-\inf\overline{f}^A(R)\big)<\frac{\varepsilon}{2},
\end{aligned}$$

and the fact that $U(f,\mathcal{P}_1)-L(f,\mathcal{P}_1)<\dfrac{\varepsilon}{2}$ shows that

$$\begin{aligned}
U(f,\mathcal{P})-\mathrm{I} &\leqslant U(f,\mathcal{P})-\mathrm{I}+U(f,\mathcal{P}_1)-U(f,\mathcal{P}_1)\\
&\leqslant U(f,\mathcal{P})-L(f,\mathcal{P}_1)+U(f,\mathcal{P}_1)-U(f,\mathcal{P}')<\varepsilon\,.
\end{aligned}$$

Similar argument can be used to show that $L(f,\mathcal{P})-\mathrm{I}>\varepsilon$. Therefore,

$$\mathrm{I}-\varepsilon<L(f,\mathcal{P})\leqslant U(f,\mathcal{P})<\mathrm{I}+\varepsilon$$

which implies that any Riemann sum of f for $\mathcal{P}$ locates in $(\mathrm{I}-\varepsilon,\mathrm{I}+\varepsilon)$. □

Notation: If $f : A \to \mathbb{R}$ is Riemann/Darboux integrable on A,

$$(\mathrm{R})\int_A f(x)\,dx = (\mathrm{D})\int_A f(x)\,dx$$

and we use $\displaystyle\int_A f(x)\,dx$ to denote this common number.

From now on, we will simply use $\bar{f}$ to denote the zero extension of f when the domain outside which the zero extension is made is clear.

6.2 The Lebesgue Theorem

In this section, we talk about another equivalent condition of Riemann/Darboux integrability, named the Lebesgue theorem. The Lebesque theorem provides a more practical way to check the Riemann/Darboux integrability in the development of theory. To understand the Lebesgue theorem, we need to talk about a new concept, sets of measure zero.

6.2.1 Volume and sets of measure zero

Definition 6.15. A bounded set $A \subseteq \mathbb{R}^n$ is said to ***have volume*** if the characteristic function or the indicator function of A, denoted by $\mathbf{1}_A$ and given by

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise}, \end{cases}$$

is Riemann integrable on A, and the number $\displaystyle\int_A \mathbf{1}_A(x)\,dx$ is called the ***volume*** of A and is denoted by $\nu(A)$. If $\nu(A) = 0$, then A is said to have volume zero or be a set of volume zero.

Remark 6.16. Not all bounded set has volume.

Proposition 6.17. *Let $A \subseteq \mathbb{R}^n$ be bounded. Then the following three statements are equivalent.*

(a) *A has volume zero;*

(b) *for every $\varepsilon > 0$, there exists finite open rectangles $S_1, \cdots, S_N$ whose sides are parallel to the coordinate axes such that*

$$A \subseteq \bigcup_{k=1}^{N} S_k \qquad \textit{and} \qquad \sum_{k=1}^{N} \nu(S_k) < \varepsilon \tag{6.2.1}$$

(c) *for every $\varepsilon > 0$, there exist finite rectangles $S_1, \cdots, S_N$ such that* (6.2.1) *holds.*

Proof. It suffices to show (a)$\Rightarrow$(b) and (c)$\Rightarrow$(a) since it is clear that (b)$\Rightarrow$(c).

"(a)$\Rightarrow$(b)" Let $\varepsilon > 0$ be given. Since A has volume zero, $\overline{\int_A} \mathbf{1}_A(x)\,dx = 0$; thus there exists a partition $\mathcal{P}$ of A such that

$$U(\mathbf{1}_A, \mathcal{P}) < \overline{\int_A} \mathbf{1}_A(x)\,dx + \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

Since $\sup\limits_{x \in \Delta} \mathbf{1}_A(x) = \begin{cases} 1 & \text{if } \Delta \cap A \neq \varnothing, \\ 0 & \text{otherwise}, \end{cases}$ we must have $\sum\limits_{\substack{\Delta \in \mathcal{P} \\ \Delta \cap A \neq \varnothing}} \nu(\Delta) < \frac{\varepsilon}{2}$. Now if $\Delta \in \mathcal{P}$ and $\Delta \cap A \neq \varnothing$, we can find an open rectangle $\square$ whose sides are parallel to the coordinate axes such that $\Delta \subseteq \square$ and $\nu(\square) < 2\nu(\Delta)$. Let $S_1, \cdots, S_N$ be those open rectangles $\square$. Then $A \subseteq \bigcup\limits_{k=1}^{N} S_k$ and $\sum\limits_{k=1}^{N} \nu(S_k) < \varepsilon$.

"(c)$\Rightarrow$(a)" Let $\varepsilon > 0$ be given. By assumption there exist rectangles $S_1, S_2, \cdots, S_N$ such that (6.2.1) holds. W.L.O.G. we can assume that the ratio of the maximum length and minimum length of sides of S_k is less than 2 for all $k = 1, \cdots, N$ (otherwise we can divide S_k into smaller rectangles so that each smaller rectangle satisfies this requirement). Then each S_k can be covered by a closed rectangle $\square_k$ whose sides are parallel to the coordinate axes with the property that $\nu(\square_k) \leqslant 2^{n-1}\sqrt{n}^{\,n}\nu(S_k)$. Let $\mathcal{P}$ be a partition of A such that for each $\Delta \in \mathcal{P}$ with $\Delta \cap A \neq \varnothing$, $\Delta \subseteq \square_k$ for some $k = 1, \cdots, N$. Then

$$U(\mathbf{1}_A, \mathcal{P}) = \sum_{\substack{\Delta \in \mathcal{P} \\ \Delta \cap A \neq \varnothing}} \nu(\Delta) \leqslant \sum_{k=1}^{N} \nu(\square_k) \leqslant 2^{n-1}\sqrt{n}^{\,n} \sum_{k=1}^{N} \nu(S_k) < 2^{n-1}\sqrt{n}^{\,n}\varepsilon;$$

thus the upper integral $\overline{\int_A} \mathbf{1}_A(x)\,dx = 0$. Since the lower integral cannot be negative, we must have $\overline{\int_A} \mathbf{1}_A(x)\,dx = \underline{\int_A} \mathbf{1}_A(x)\,dx = 0$ which shows that A has volume zero. $\square$

Example 6.18. Each point in $\mathbb{R}^n$ has volume zero.

Definition 6.19. A set $A \subseteq \mathbb{R}^n$ (not necessarily bounded) is said to ***have measure zero*** (測度為零) or be ***a set of measure zero*** (零測度集) if for every $\varepsilon > 0$, there exist countable many rectangles $S_1, S_2, \cdots$ such that $\{S_k\}_{k=1}^{\infty}$ is a cover of A $\big($that is, $A \subseteq \bigcup\limits_{k=1}^{\infty} S_k\big)$ and $\sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon$.

Example 6.20. The real line $\mathbb{R}\times\{0\}$ on $\mathbb{R}^2$ has measure zero: for any given $\varepsilon>0$, let $S_k=[-k,k]\times\big[\frac{-\varepsilon}{2^{k+3}k},\frac{\varepsilon}{2^{k+3}k}\big]$. Then

$$\mathbb{R}\times\{0\}\subseteq\bigcup_{k=1}^{\infty}S_k \quad\text{and}\quad \sum_{k=1}^{\infty}\nu(S_k)=\sum_{k=1}^{\infty}2k\cdot\frac{2\varepsilon}{2^{k+3}k}=\sum_{k=1}^{\infty}\frac{\varepsilon}{2^{k+1}}=\frac{\varepsilon}{2}<\varepsilon\,.$$

Similarly, any hyperplane in $\mathbb{R}^n$ also has measure zero.

Proposition 6.21. *Let $A\subseteq\mathbb{R}^n$ be a set of measure zero. If $B\subseteq A$, then B also has measure zero.*

Modifying the proof of Proposition 6.17, we can also conclude the following

Proposition 6.22. *A set $A\subseteq\mathbb{R}^n$ has measure zero if and only if for every $\varepsilon>0$, there exist countable many open rectangles S_1, S_2, $\cdots$ whose sides are parallel to the coordinate axes such that $A\subseteq\bigcup\limits_{k=1}^{\infty}S_k$ and $\sum\limits_{k=1}^{\infty}\nu(S_k)<\varepsilon$.*

Remark 6.23. If a set A has volume zero, then it has measure zero.

Proposition 6.24. *Let $K\subseteq\mathbb{R}^n$ be a compact set of measure zero. Then K has volume zero.*

Proof. Let $\varepsilon>0$ be given. Then there are countable open rectangles $S_1,S_2,\cdots$ such that

$$K\subseteq\bigcup_{k=1}^{\infty}S_k \qquad\text{and}\qquad \sum_{k=1}^{\infty}\nu(S_k)<\varepsilon\,.$$

Since $\{S_k\}_{k=1}^{\infty}$ is an open cover of K, by the compactness of K there exists $N>0$ such that $K\subseteq\bigcup\limits_{k=1}^{N}S_k$, while $\sum\limits_{k=1}^{N}\nu(S_k)\leqslant\sum\limits_{k=1}^{\infty}\nu(S_k)<\varepsilon$. As a consequence, K has volume zero. ▫

Since the boundary of a rectangle has measure zero, we also have the following

Corollary 6.25. *Let $S\subseteq\mathbb{R}^n$ be a bounded rectangle with positive volume. Then S is not a set of measure zero.*

Theorem 6.26. *If A_1, A_2, $\cdots$ are sets of measure zero in $\mathbb{R}^n$, then $\bigcup\limits_{k=1}^{\infty}A_k$ has measure zero.*

Proof. Let $\varepsilon > 0$ be given. Since $A_k's$ are sets of measure zero, there exist countable rectangles $\{S_j^{(k)}\}_{j=1}^{\infty}$, such that

$$A_k \subseteq \bigcup_{j=1}^{\infty} S_j^{(k)} \quad \text{and} \quad \sum_{j=1}^{\infty} \nu(S_j^{(k)}) < \frac{\varepsilon}{2^{k+1}} \qquad \forall\, k \in \mathbb{N}.$$

Consider the collection consisting of all $S_j^{(k)}$'s. Since there are countable many rectangles in this collection, we can label them as S_1, S_2, $\cdots$, and we have

$$\bigcup_{k=1}^{\infty} A_k \subseteq \bigcup_{k=1}^{\infty}\bigcup_{j=1}^{\infty} S_j^{(k)} = \bigcup_{\ell=1}^{\infty} S_\ell$$

and

$$\sum_{k=1}^{\infty} \nu(S_\ell) = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \nu(S_j^{(k)}) \leqslant \sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2} < \varepsilon .$$

Therefore, $\bigcup\limits_{k=1}^{\infty} A_k$ has measure zero. □

Corollary 6.27. *The set of rational numbers in $\mathbb{R}$ has measure zero.*

Theorem 6.28. *Let $A \subseteq \mathbb{R}^n$ be bounded and $B \subseteq \mathbb{R}^m$ be a set of measure zero. Then $A \times B$ has measure zero in $\mathbb{R}^{n+m}$.*

Proof. Let $\varepsilon > 0$ be given. Since A is bounded, there exist a bounded rectangle R such that $A \subseteq R$. Since B has measure zero, there exist countable rectangles $\{S_k\}_{k=1}^{\infty} \subseteq \mathbb{R}^m$ such that

$$B \subseteq \bigcup_{k=1}^{\infty} S_k \quad \text{and} \quad \sum_{k=1}^{\infty} \nu_m(S_k) < \frac{\varepsilon}{\nu(R)} .$$

Then $A \times B \subseteq \bigcup\limits_{k=1}^{\infty} (R \times S_k)$, and

$$\sum_{k=1}^{\infty} \nu_{n+m}(R \times S_k) = \sum_{k=1}^{\infty} \nu_n(R)\nu_m(S_k) = \nu_n(R) \sum_{k=1}^{\infty} \nu_m(S_k) < \varepsilon .$$

Since $R \times S_k$ is a rectangle for all $k \in \mathbb{N}$, we conclude that $A \times B$ has measure zero. □

6.2.2 The Lebesgue theorem

在之前我們提到了函數 Riemann 可積的兩個等價條件：Riemann's condition 和 Darboux 定理。在這一節中，我們將引進函數是 Riemman 可積的另一個等價條件。這個等價條件

說的是一個函數 f 在 A 上是 Riemann 可積的若且唯若 f 的延拓 $\overline{f}^A$（在函數可積分的定義中有定義）的不連續點所構成的集合其測度為零。為了證明這個敘述，我們先對一個函數的連續點做一個量化的刻劃。這個刻劃的方式，可以很容易用來檢驗一個函數在一個點是否連續。

Definition 6.29. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. For any $x \in \mathbb{R}^n$, the ***oscillation*** of f at x is the quantity

$$\operatorname{osc}(f, x) \equiv \inf_{\delta>0} \sup_{x_1,x_2 \in B(x,\delta)} \big|f(x_1) - f(x_2)\big|.$$

我們注意到在上述定義中被取 infimum 的這個量 $h(\delta; x) \equiv \sup\limits_{x_1,x_2 \in B(x,\delta)} \big|f(x_1) - f(x_2)\big|$ 是個 δ 的遞減函數（x 固定），而 $\operatorname{osc}(f, x)$ 則是 $h(\delta; x)$ 當 $\delta \to 0$ 時的極限。另外，我們也注意到說 $h(\delta; x)$ 也可以寫成 $\sup\limits_{y \in B(x,\delta)} f(y) - \inf\limits_{y \in B(x,\delta)} f(y)$.

以下的 Lemma 是關於如何檢驗一個函數在一個點是連續的。

Lemma 6.30. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, and $x_0 \in \mathbb{R}^n$. Then f is continuous at x_0 if and only if* $\operatorname{osc}(f, x_0) = 0$.

Proof. "$\Rightarrow$" Let $\varepsilon > 0$ be given. Since f is continuous at x_0,

$$\exists\, \delta > 0 \ni \big|f(x) - f(x_0)\big| < \frac{\varepsilon}{3} \qquad \text{whenever } x \in B(x_0, \delta).$$

In particular, for any $x_1, x_2 \in B(x_0, \delta)$,

$$\big|f(x_1) - f(x_2)\big| \leqslant \big|f(x_1) - f(x_0)\big| + \big|f(x_0) - f(x_2)\big| < \frac{2\varepsilon}{3};$$

thus $\sup\limits_{x_1,x_2 \in B(x_0,\delta)} \big|f(x_1) - f(x_2)\big| \leqslant \dfrac{2\varepsilon}{3}$ which further implies that

$$0 \leqslant \operatorname{osc}(f, x_0) \leqslant \frac{2\varepsilon}{3} < \varepsilon.$$

Since ε is given arbitrarily, $\operatorname{osc}(f, x_0) = 0$.

"$\Leftarrow$" Let $\varepsilon > 0$ be given. By the definition of infimum, there exists $\delta > 0$ such that

$$\sup_{x_1,x_2 \in B(x_0,\delta)} \big|f(x_1) - f(x_2)\big| < \varepsilon.$$

In particular, $\big|f(x) - f(x_0)\big| \leqslant \sup\limits_{x_1,x_2 \in B(x_0,\delta)} \big|f(x_1) - f(x_2)\big| < \varepsilon$ for all $x \in B(x_0, \delta)$. $\square$

Lemma 6.31. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Then for all $\varepsilon > 0$, the set $D_\varepsilon = \big\{x \in \mathbb{R}^n \,\big|\, \operatorname{osc}(f,x) \geqslant \varepsilon\big\}$ is closed.*

Proof. Suppose that $\{y_k\}_{k=1}^\infty \subseteq D_\varepsilon$ and $y_k \to y$. Then for any $\delta > 0$, there exists $N > 0$ such that $y_k \in B(y,\delta)$ for all $k \geqslant N$. Since $B(y,\delta)$ is open, for each $k \geqslant N$ there exists $\delta_k > 0$ such that $B(y_k,\delta_k) \subseteq B(y,\delta)$; thus we find that

$$\sup_{x_1,x_2 \in B(y_k,\delta_k)} \big|f(x_1) - f(x_2)\big| \leqslant \sup_{x_1,x_2 \in B(y,\delta)} \big|f(x_1) - f(x_2)\big| \qquad \forall\, k \geqslant N\,.$$

The inequality above implies that $\operatorname{osc}(f,y) \geqslant \varepsilon$; thus $y \in D_\varepsilon$ and D_ε is closed. □

Theorem 6.32 (Lebesgue)**.** *Let $A \subseteq \mathbb{R}^n$ be a bounded set, $f : A \to \mathbb{R}$ be a bounded function, and $\overline{f}^A$ be the extension of f by zero outside A; that is,*

$$\overline{f}^A(x) = \begin{cases} f(x) & \textit{if } x \in A\,, \\ 0 & \textit{otherwise}\,. \end{cases}$$

Then f is Riemann integrable on A if and only if the collection of discontinuity of $\overline{f}^A$ is a set of measure zero.

Proof. Let $D = \big\{x \in \mathbb{R}^n \,\big|\, \operatorname{osc}(\overline{f}^A, x) > 0\big\}$ and $D_\varepsilon = \big\{x \in \mathbb{R}^n \,\big|\, \operatorname{osc}(\overline{f}^A, x) \geqslant \varepsilon\big\}$. We remark here that $D = \bigcup\limits_{k=1}^\infty D_{\frac{1}{k}}$.

"$\Rightarrow$" We show that $D_{\frac{1}{k}}$ has measure zero for all $k \in \mathbb{N}$ (if so, then Theorem 6.26 implies that D has measure zero).

Let $k \in \mathbb{N}$ be fixed, and $\varepsilon > 0$ be given. By Riemann's condition there exists a partition $\mathcal{P}$ of A such that

$$\sum_{\Delta \in \mathcal{P}} \Big[\sup_{x \in \Delta} \overline{f}^A(x) - \inf_{x \in \Delta} \overline{f}^A(x)\Big] \nu(\Delta) < \frac{\varepsilon}{k}\,.$$

Define

$$D^{(1)}_{\frac{1}{k}} \equiv \big\{x \in D_{\frac{1}{k}} \,\big|\, x \in \partial\Delta \text{ for some } \Delta \in \mathcal{P}\big\}\,,$$
$$D^{(2)}_{\frac{1}{k}} \equiv \big\{x \in D_{\frac{1}{k}} \,\big|\, x \in \operatorname{int}(\Delta) \text{ for some } \Delta \in \mathcal{P}\big\}\,.$$

Then $D_{\frac{1}{k}} = D^{(1)}_{\frac{1}{k}} \cup D^{(2)}_{\frac{1}{k}}$. We note that $D^{(1)}_{\frac{1}{k}}$ has measure zero since it is contained in $\bigcup\limits_{\Delta \in \mathcal{P}} \partial\Delta$ while each $\partial\Delta$ has measure zero. Now we show that $D^{(2)}_{\frac{1}{k}}$ also has measure zero. Let $C = \big\{\Delta \in \mathcal{P} \,\big|\, \text{int}(\Delta) \cap D_{\frac{1}{k}} \neq \varnothing\big\}$. Then $D^{(2)}_{\frac{1}{k}} \subseteq \bigcup\limits_{\Delta \in C} \Delta$. Moreover, we also note that if $\Delta \in C$, $\sup\limits_{x\in\Delta} \overline{f}^A(x) - \inf\limits_{x\in\Delta} \overline{f}^A(x) \geqslant \dfrac{1}{k}$. In fact, if $\Delta \in C$, there exists $y \in \text{int}(\Delta) \cap D_{\frac{1}{k}}$; thus choosing $\delta > 0$ such that $B(y,\delta) \subseteq \text{int}(\Delta)$,

$$\begin{aligned}\sup_{x\in\Delta} \overline{f}^A(x) - \inf_{x\in\Delta} \overline{f}^A(x) &= \sup_{x_1,x_2\in\Delta} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| \geqslant \sup_{x_1,x_2\in B(y,\delta)} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| \\ &\geqslant \inf_{\delta>0} \sup_{x_1,x_2\in B(y,\delta)} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| = \text{osc}(\overline{f}^A, y) \geqslant \frac{1}{k}.\end{aligned}$$

As a consequence,

$$\frac{1}{k}\sum_{\Delta\in C} \nu(\Delta) \leqslant \sum_{\Delta\in\mathcal{P}} \Big[\sup_{x\in\Delta} \overline{f}^A(x) - \inf_{x\in\Delta} \overline{f}^A(x)\Big]\nu(\Delta) = U(f,\mathcal{P}) - L(f,\mathcal{P}) < \frac{\varepsilon}{k}$$

which implies that $\sum\limits_{\Delta\in C} \nu(\Delta) < \varepsilon$. In other words, we establish that $D^{(2)}_{\frac{1}{k}}$ has measure zero. Therefore, $D_{\frac{1}{k}}$ has measure zero for all $k \in \mathbb{N}$; thus D has measure zero.

"$\Leftarrow$" Let R be a bounded closed rectangle with sides parallel to the coordinate axes and $\bar{A} \subseteq \text{int}(R)$, and $\varepsilon > 0$ be given. Define $\varepsilon' = \dfrac{\varepsilon}{2\|f\|_\infty + \nu(R) + 1}$, where $\|f\|_\infty = \sup\limits_{x\in A} |f(x)|$.

1. Since $D_{\varepsilon'}$ is a subset of D, Proposition 6.21 implies that $D_{\varepsilon'}$ has measure zero; thus Proposition 6.22 provides open rectangles S_1, S_2, $\cdots$ whose sides are parallel to the coordinate axes such that $D_{\varepsilon'} \subseteq \bigcup\limits_{k=1}^{\infty} S_k$, and $\sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon'$. In addition, we can assume that $S_k \subseteq R$ for all $k \in \mathbb{N}$ since $D_{\varepsilon'} \subseteq R$.

2. Since $D_{\varepsilon'} \subseteq R$ is bounded, Lemma 6.31 implies that $D_{\varepsilon'}$ is compact; thus $D_{\varepsilon'} \subseteq \bigcup\limits_{k=1}^{N} S_k$ for some $N \in \mathbb{N}$.

Let $\square_k = \overline{S_k}$, and $\mathcal{P}_1$ be a partition of R satisfying

(a) For each $\Delta \in \mathcal{P}_1$ with $\Delta \cap D_{\varepsilon'} \neq \varnothing$, $\Delta \subseteq \square_k$ for some $k = 1, \cdots, N$.

(b) For each $k = 1, \cdots, N$, $\square_k$ is the union of rectangles in $\mathcal{P}_1$.

(c) Some collection of $\Delta \in \mathcal{P}_1$ forms a partition $\mathcal{P}_2$ of A.

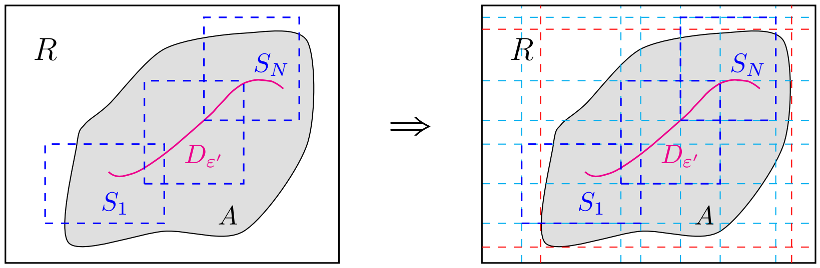


Figure 6.2: Constructing partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ from finite rectangles $S_1, \cdots, S_N$

Rectangles in $\mathcal{P}_1$ fall into two families: $C_1 = \{\Delta \in \mathcal{P}_1 \,|\, \Delta \subseteq \square_k \text{ for some } k = 1, \cdots, N\}$, and $C_2 = \{\Delta \in \mathcal{P}_1 \,|\, \Delta \nsubseteq \square_k \text{ for all } k = 1, \cdots, N\}$. By the definition of the oscillation function, for $x \notin D_{\varepsilon'}$ we let $\delta_x > 0$ be such that

$$\sup_{x \in B(y,\delta_y)} \overline{f}^A(y) - \inf_{x \in B(y,\delta_y)} \overline{f}^A(y) = \sup_{x_1,x_2 \in B(x,\delta_x)} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| < \varepsilon' .$$

Since $K = \bigcup\limits_{\Delta \in C_2} \Delta$ is compact, there exists $r > 0$ (the Lebesgue number associated with K and open cover $\bigcup\limits_{x \in K} B(x, \delta_x)$) such that for each $a \in K$, $B(a, r) \subseteq B(y, \delta_y)$ for some $y \in K$. Let $\mathcal{P}'$ be a refinement of $\mathcal{P}_1$ such that $\|\mathcal{P}'\| < r$. Then if $\Delta' \in \mathcal{P}'$ satisfies that $\Delta' \subseteq \Delta$ for some $\Delta \in C_2$, we must have $\Delta' \subseteq B(y, \delta_y)$ for some $y \in K$; thus

$$\begin{aligned}\sup_{x \in \Delta'} \overline{f}^A(x) - \inf_{x \in \Delta'} \overline{f}^A(x) &\leqslant \sup_{x \in B(y,\delta_y)} \overline{f}^A(y) - \inf_{x \in B(y,\delta_y)} \overline{f}^A(y) \\ &= \sup_{x_1,x_2 \in B(y,\delta_y)} \big|\overline{f}^A(x_1) - \overline{f}^A(x_2)\big| < \varepsilon'\end{aligned}$$

if $\Delta' \subseteq \Delta$ for some $\Delta \in C_2$. Let $\mathcal{P} = \{\Delta' \in \mathcal{P}' \,|\, \Delta' \subseteq \Delta \text{ for some } \Delta \in \mathcal{P}_2\}$. Then $\mathcal{P}$ is a partition of A and

$$\begin{aligned}U(f, \mathcal{P}) - L(f, \mathcal{P}) &= \Big(\sum_{\substack{\Delta' \in \mathcal{P}' \\ \Delta' \subseteq \Delta \in C_1}} + \sum_{\substack{\Delta' \in \mathcal{P}' \\ \Delta' \subseteq \Delta \in C_2}} \Big) \big(\sup_{x \in \Delta'} \overline{f}^A(x) - \inf_{x \in \Delta'} \overline{f}^A(x) \big) \nu(\Delta') \\ &\leqslant 2\|f\|_\infty \sum_{\substack{\Delta' \in \mathcal{P}' \\ \Delta' \subseteq \Delta \in C_1}} \nu(\Delta') + \varepsilon' \sum_{\substack{\Delta' \in \mathcal{P}' \\ \Delta' \subseteq \Delta \in C_2}} \nu(\Delta') \\ &\leqslant 2\|f\|_\infty \sum_{\Delta \in \mathcal{P} \cap C_1} \nu(\Delta) + \varepsilon' \nu(R) \\ &\leqslant 2\|f\|_\infty \sum_{k=1}^{N} \nu(S_k) + \varepsilon' \nu(R) < \big(2\|f\|_\infty + \nu(R)\big)\varepsilon' \leqslant \varepsilon;\end{aligned}$$

thus f is Riemann integrable on A by Riemann's condition. □

Example 6.33. Let $A = \mathbb{Q} \cap [0,1]$, and $f : A \to \mathbb{R}$ be the constant function $f \equiv 1$. Then

$$\bar{f}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0,1], \\ 0 & \text{otherwise}. \end{cases}$$

The collection of points of discontinuity of $\bar{f}$ is $[0,1]$ which, by Corollary 6.25, cannot be a set of measure zero; thus f is not Riemann integrable.

Another way to see that f is not Riemann integrable is $U(f,\mathcal{P}) = 1$ and $L(f,\mathcal{P}) = 0$ for all partitions $\mathcal{P}$ of A.

Corollary 6.34. *A bounded set $A \subseteq \mathbb{R}^n$ has volume if and only if the boundary of A has measure zero.*

Proof. It suffices to show that the collection of discontinuities of the function $\mathbf{1}_A$ (which is the same as $\overline{\mathbf{1}_A}^A$) is indeed ∂A.

1. If $x_0 \notin \partial A$, then there exists $\delta > 0$ such that either $B(x_0,\delta) \subseteq A$ or $B(x_0,\delta) \subseteq A^\complement$; thus $\mathbf{1}_A$ is continuous at $x_0 \notin \partial A$ since $\mathbf{1}_A(x)$ is constant for all $x \in B(x_0,\delta)$.

2. On the other hand, if $x_0 \in \partial A$, then there exists $x_k \in A$, $y_k \in A^\complement$ such that $x_k \to x_0$ and $y_k \to x_0$ as $k \to \infty$. This implies that $\mathbf{1}_A$ cannot be continuous at x_0 since $\mathbf{1}_A(x_k) = 1$ while $\mathbf{1}_A(y_k) = 0$ for all $k \in \mathbb{N}$.

As a consequence, the collection of discontinuity of $\mathbf{1}_A$ is exactly ∂A, and the corollary follows from Lebesgue's theorem. □

Corollary 6.35. *Let $A \subseteq \mathbb{R}^n$ be a bounded set with volume. A bounded function $f : A \to \mathbb{R}$ is Riemann integrable on A if and only if the collection of discontinuities of f has measure zero. In particular, a bounded function $f : A \to \mathbb{R}$ with a finite or countable number of points of discontinuity is Riemann integrable on A.*

Proof. Note that

$$\{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\bar{f}, x) > 0\} \subseteq \partial A \cup \{x \in A \,|\, f \text{ is discontinuous at } x\}$$

and

$$\{x \in A \,|\, f \text{ is discontinuous at } x\} \subseteq \{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\bar{f}, x) > 0\};$$

thus by the fact that ∂A has measure zero (Corollary 6.34) and Theorem 6.26 we conclude that the collection of discontinuities of $\bar{f}$ has measure zero if and only if the collection of discontinuities of f has measure zero. □

Corollary 6.36. *A bounded function is integrable on a compact set of measure zero.*

Proof. If $f : K \to \mathbb{R}$ is bounded, and K is a compact set of measure zero, then the collection of discontinuities of $\bar{f}$ is a subset of K. ◻

Corollary 6.37. *Suppose that $A, B \subseteq \mathbb{R}^n$ are bounded sets with volume, and $f : A \to \mathbb{R}$ is Riemann integrable on A. Then f is Riemann integrable on $A \cap B$.*

Proof. By the inclusion

$$\big\{x \in \operatorname{int}(A \cap B) \,\big|\, \operatorname{osc}(\overline{f}^{A\cap B}, x) > 0\big\} \subseteq \big\{x \in \mathbb{R}^n \,\big|\, \operatorname{osc}(\overline{f}^{A}, x) > 0\big\},$$

we find that

$$\begin{aligned}\big\{x \in \mathbb{R}^n \,\big|\, \operatorname{osc}(\overline{f}^{A\cap B}, x) > 0\big\} &\subseteq \partial(A \cap B) \cup \big\{x \in \operatorname{int}(A \cap B) \,\big|\, \operatorname{osc}(\overline{f}^{A\cap B}, x) > 0\big\} \\ &\subseteq \partial A \cup \partial B \cup \big\{x \in \mathbb{R}^n \,\big|\, \operatorname{osc}(\overline{f}^{A}, x) > 0\big\}.\end{aligned}$$

Since ∂A and ∂B both have measure zero, the integrability of f on $A \cap B$ then follows from the integrability of f on A and the Lebesgue Theorem. ◻

Remark 6.38. Suppose that $A \subseteq \mathbb{R}^n$ is a bounded set of measure zero. Even if $f : A \to \mathbb{R}$ is continuous, f might not be Riemann integrable. For example, the function f given in Example 6.33 is not Riemann integrable even though f is continuous on A.

Remark 6.39. When $f : A \to \mathbb{R}$ is Riemann integrable on A, it is not necessary that A has volume. For example, the zero function is Riemann integrable on $A = \mathbb{Q} \cap [0, 1]$ even though A does not has volume.

Corollary 6.40 (Lebesgue's Differentiation Theorem, a simple version)**.** *Let $A \subseteq \mathbb{R}^n$ be a bounded set with volume, and $f : A \to \mathbb{R}$ be bounded and Riemann integrable on A. Then*

$$\lim_{r \to 0} \frac{1}{\nu(B(x_0, r) \cap A)} \int_{B(x_0,r)\cap A} f(x)\, dx = f(x_0) \tag{6.2.2}$$

for almost every $x_0 \in A$; that is, the equality above does not hold only for x_0 from a set of measure zero.

Proof. Let $\varepsilon > 0$ be given, and suppose that $\bar{f}$, the zero extension of f outside A, is continuous at x_0. Then there exists $\delta > 0$ such that

$$\big|\bar{f}(x) - \bar{f}(x_0)\big| < \frac{\varepsilon}{2} \qquad \forall\, x \in B(x_0, \delta) \cap A.$$

Since ∂A has measure zero, by the fact that $\partial(B(x_0,r)\cap A)\subseteq \partial B(x_0,r)\cup\partial A$ we find that $\partial(B(x_0,r)\cap A)$ also has measure zero for all $r>0$. In other words, $B(x_0,r)\cap A$ has volume. Then if $0<r<\delta$,

$$\begin{aligned}
&\Big|\frac{1}{\nu(B(x_0,r)\cap A)}\int_{B(x_0,r)\cap A} f(x)\,dx - f(x_0)\Big| \\
&\quad = \Big|\frac{1}{\nu(B(x_0,r)\cap A)}\int_{B(x_0,r)\cap A}\big(\bar{f}(x)-\bar{f}(x_0)\big)\,dx\Big| \\
&\quad \leqslant \frac{1}{\nu(B(x_0,r)\cap A)}\int_{B(x_0,r)\cap A}\big|\bar{f}(x)-\bar{f}(x_0)\big|\,dx \\
&\quad \leqslant \frac{\varepsilon}{2}\frac{1}{\nu(B(x_0,r)\cap A)}\int_{B(x_0,r)\cap A} 1\,dx = \frac{\varepsilon}{2} < \varepsilon.
\end{aligned}$$

This implies that (6.2.2) holds for all x_0 at which $\bar{f}$ is continuous. The theorem then follows from the Lebesgue theorem. ▫

6.3 Properties of the Integrals

Proposition 6.41. *Let $A\subseteq\mathbb{R}^n$ be a bounded set, and $f,g:A\to\mathbb{R}$ be bounded functions. Then*

(a) *If $B\subseteq A$, then $\underline{\int}_A (f\mathbf{1}_B)(x)\,dx = \underline{\int}_B f(x)\,dx$ and $\overline{\int}_A (f\mathbf{1}_B)(x)\,dx = \overline{\int}_B f(x)\,dx$.*

(b) $\underline{\int}_A f(x)\,dx + \underline{\int}_A g(x)\,dx \leqslant \underline{\int}_A (f+g)(x)\,dx \leqslant \overline{\int}_A (f+g)(x)\,dx \leqslant \overline{\int}_A f(x)\,dx + \overline{\int}_A g(x)\,dx.$

(c) *If $c\geqslant 0$, then $\underline{\int}_A (cf)(x)\,dx = c\underline{\int}_A f(x)\,dx$ and $\overline{\int}_A (cf)(x)\,dx = c\overline{\int}_A f(x)\,dx$. If $c<0$, then $\underline{\int}_A (cf)(x)\,dx = c\overline{\int}_A f(x)\,dx$ and $\overline{\int}_A (cf)(x)\,dx = c\underline{\int}_A f(x)\,dx$.*

(d) *If $f\leqslant g$ on A, then $\underline{\int}_A f(x)\,dx \leqslant \underline{\int}_A g(x)\,dx$ and $\overline{\int}_A f(x)\,dx \leqslant \overline{\int}_A g(x)\,dx$.*

(e) *If A has volume zero, then f is Riemann integrable on A, and $\int_A f(x)\,dx = 0$.*

Proof. We only prove (a), (b), (c) and (e) since (d) is trivial.

(a) Let $\varepsilon > 0$ be given. By the definition of the lower integral, there exist partition $\mathcal{P}_A$ of A and $\mathcal{P}_B$ of B such that

$$\underline{\int}_A (f\mathbf{1}_B)(x)\,dx - \varepsilon < L(f\mathbf{1}_B, \mathcal{P}_A) = \sum_{\Delta\in\mathcal{P}_A} \inf_{x\in\Delta} \overline{f\mathbf{1}_B}^A(x)\nu(\Delta)$$

and

$$\underline{\int}_B f(x)\,dx - \frac{\varepsilon}{2} < L(f, \mathcal{P}_B) = \sum_{\Delta\in\mathcal{P}_B} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta).$$

Let $\mathcal{P}'_A$ be a refinement of $\mathcal{P}_A$ such that some collection of rectangles in $\mathcal{P}'_A$ forms a partition of B. Denote this partition of B by $\mathcal{P}'_B$. Since $\inf\limits_{x\in\Delta} \overline{f}^B(x) \leqslant 0$ if $\Delta \in \mathcal{P}'_A \backslash \mathcal{P}'_B$, Proposition 6.10 implies that

$$\begin{aligned}
\underline{\int}_A (f\mathbf{1}_B)(x)\,dx - \varepsilon < L(f\mathbf{1}_B, \mathcal{P}_A) \leqslant L(f\mathbf{1}_B, \mathcal{P}'_A) &= \sum_{\Delta\in\mathcal{P}'_A} \inf_{x\in\Delta} \overline{f\mathbf{1}_B}^A(x)\nu(\Delta) \\
&= \Big(\sum_{\Delta\in\mathcal{P}'_A\backslash\mathcal{P}'_B} + \sum_{\Delta\in\mathcal{P}'_B} \Big) \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) \\
&\leqslant \sum_{\Delta\in\mathcal{P}'_B} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) = L(f, \mathcal{P}'_B) \leqslant \underline{\int}_B f(x)\,dx.
\end{aligned}$$

On the other hand, let $\widetilde{\mathcal{P}}_A$ be a partition of A such that $\mathcal{P}_B \subseteq \widetilde{\mathcal{P}}_A$ and

$$\sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} \nu(\Delta) \leqslant \frac{\varepsilon}{2(M+1)},$$

where $M > 0$ is an upper bound of $|f|$. Then

$$\sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) \geqslant -M \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} \nu(\Delta) \geqslant -\frac{\varepsilon}{2}$$

which further implies that

$$\begin{aligned}
\underline{\int}_A (f\mathbf{1}_B)(x)\,dx \geqslant L(f\mathbf{1}_B, \widetilde{\mathcal{P}}_A) &= \sum_{\Delta\in\widetilde{\mathcal{P}}_A} \inf_{x\in\Delta} \overline{f\mathbf{1}_B}^A(x)\nu(\Delta) \\
&= \Big(\sum_{\Delta\in\mathcal{P}_B} + \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} + \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B=\varnothing} \Big) \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) \\
&= L(f, \mathcal{P}_B) + \sum_{\Delta\in\widetilde{\mathcal{P}}_A\backslash\mathcal{P}_B,\, \Delta\cap B\neq\varnothing} \inf_{x\in\Delta} \overline{f}^B(x)\nu(\Delta) > \underline{\int}_B f(x)\,dx - \varepsilon.
\end{aligned}$$

Therefore, we establish that

$$\underline{\int_B} f(x)\,dx - \varepsilon < \underline{\int_A} (f\mathbf{1}_B)(x)\,dx < \underline{\int_B} f(x)\,dx + \varepsilon\,.$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that $\underline{\int_A} (f\mathbf{1}_B)(x)\,dx = \underline{\int_B} f(x)\,dx$. Similar argument can be applied to conclude that $\overline{\int_A} (f\mathbf{1}_B)(x)\,dx = \overline{\int_B} f(x)\,dx$.

(b) Let $\varepsilon > 0$ be given. By the definition of the lower integral, there exist partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of A such that

$$\underline{\int_A} f(x)\,dx - \frac{\varepsilon}{2} < L(f,\mathcal{P}_1) \quad \text{and} \quad \underline{\int_A} g(x)\,dx - \frac{\varepsilon}{2} < L(g,\mathcal{P}_2)\,.$$

Let $\mathcal{P}$ be a common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Then

$$\begin{aligned}
\underline{\int_A} f(x)\,dx + \underline{\int_A} g(x)\,dx - \varepsilon < L(f,\mathcal{P}_1) + L(f,\mathcal{P}_2) &\leqslant L(f,\mathcal{P}) + L(g,\mathcal{P}) \\
&= \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta} \bar{f}(x)\nu(\Delta) + \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta} \bar{g}(x)\nu(\Delta) \\
&\leqslant \sum_{\Delta\in\mathcal{P}} \inf_{x\in\Delta} (\bar{f}+\bar{g})(x)\nu(\Delta) = L(f+g,\mathcal{P}) \leqslant \underline{\int_A} (f+g)(x)\,dx\,.
\end{aligned}$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that

$$\underline{\int_A} f(x)\,dx + \underline{\int_A} g(x)\,dx \leqslant \underline{\int_A} (f+g)(x)\,dx\,.$$

Similarly, we have $\overline{\int_A} (f+g)(x)\,dx \leqslant \overline{\int_A} f(x)\,dx + \overline{\int_A} g(x)\,dx$; thus (b) is established.

(c) It suffices to show the case $c = -1$. Let $\varepsilon > 0$ be given. Then there exist partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of A such that

$$\underline{\int_A} -f(x)\,dx - \varepsilon < L(-f,\mathcal{P}_1) \qquad \text{and} \qquad U(f,\mathcal{P}_2) < \overline{\int_A} f(x)\,dx + \varepsilon\,.$$

Let $\mathcal{P}$ be the common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. Then

$$\underline{\int_A} -f(x)\,dx - \varepsilon < L(-f,\mathcal{P}_1) \leqslant L(-f,\mathcal{P}) \leqslant \underline{\int_A} -f(x)\,dx$$

and

$$\overline{\int_A} f(x)\,dx \leqslant U(f,\mathcal{P}) \leqslant U(f,\mathcal{P}_2) < \overline{\int_A} f(x)\,dx + \varepsilon\,.$$

By the fact that

$$L(-f,\mathcal{P})=\sum_{\Delta\in\mathcal{P}}\inf_{x\in\Delta}\overline{(-f)}^A(x)\nu(\Delta)=-\sum_{\Delta\in\mathcal{P}}\sup_{x\in\Delta}\overline{f}^A(x)\nu(\Delta)=-U(f,\mathcal{P}),$$

we find that

$$\underline{\int_A}-f(x)\,dx-\varepsilon<L(-f,\mathcal{P})=-U(f,\mathcal{P})\leqslant-\overline{\int_A}f(x)\,dx$$

and

$$\underline{\int_A}-f(x)\,dx\geqslant L(-f,\mathcal{P})=-U(f,\mathcal{P})>-\overline{\int_A}f(x)\,dx-\varepsilon.$$

Therefore,

$$\underline{\int_A}-f(x)\,dx-\varepsilon<-\overline{\int_A}f(x)\,dx<\underline{\int_A}-f(x)\,dx+\varepsilon.$$

Since $\varepsilon>0$ is given arbitrarily, we conclude (c).

(e) Since f is bounded on A, there exist $M>0$ such that $-M\leqslant f(x)\leqslant M$ for all $x\in A$. Therefore, $-1_A\leqslant\dfrac{f}{M}\leqslant 1_A$ on A; thus (c) and (d) imply that

$$0=\int_A\mathbf{1}_A(x)\,dx=\overline{\int_A}\mathbf{1}_A(x)\,dx\geqslant\overline{\int_A}\frac{f(x)}{M}\,dx=\frac{1}{M}\overline{\int_A}f(x)\,dx$$

which implies that $\overline{\int_A}f(x)\,dx\leqslant 0$. Similarly, $\overline{\int_A}-f(x)\,dx\leqslant 0$ which further implies that $\underline{\int_A}f(x)\,dx\geqslant 0$. Therefore, by Corollary 6.12 we conclude that

$$0\leqslant\underline{\int_A}f(x)\,dx\leqslant\overline{\int_A}f(x)\,dx\leqslant 0$$

which implies that f is Riemann integrable on A and $\int_A f(x)\,dx=0$. □

Remark 6.42. Let $A\subseteq\mathbb{R}^n$ be a bounded set.

1. If $f:A\to\mathbb{R}$ is a bounded function, then (a) of Proposition 6.41 shows that if $B\subseteq A$, then f is Riemann integrable on B if and only if $f\mathbf{1}_B$ is Riemann integrable on A.

2. If $f,g:A\to\mathbb{R}$ are bounded functions, then (b) of Proposition 6.41 also implies that

$$\underline{\int_A}(f-g)(x)\,dx\leqslant\underline{\int_A}f(x)\,dx-\underline{\int_A}g(x)\,dx\ \text{ and }\ \overline{\int_A}f(x)\,dx-\overline{\int_A}g(x)\,dx\leqslant\overline{\int_A}(f-g)(x)\,dx.$$

Corollary 6.43. *Let $A, B \subseteq \mathbb{R}^n$ be bounded sets such that $A \cap B$ has volume zero, and $f : A \cup B \to \mathbb{R}$ be a bounded function. Then*

$$\underline{\int}_A f(x)\,dx + \underline{\int}_B f(x)\,dx \leqslant \underline{\int}_{A\cup B} f(x)\,dx \leqslant \overline{\int}_{A\cup B} f(x)\,dx \leqslant \overline{\int}_A f(x)\,dx + \overline{\int}_B f(x)\,dx\,.$$

Proof. Note that $f\mathbf{1}_A + f\mathbf{1}_B = f\mathbf{1}_{A\cup B} + f\mathbf{1}_{A\cap B}$ on $A \cup B$. Therefore, (a), (b) of Proposition 6.41 and Remark 6.42 implies that

$$\begin{aligned}\underline{\int}_A f(x)\,dx + \underline{\int}_B f(x)\,dx &= \underline{\int}_{A\cup B} (f\mathbf{1}_A)(x)\,dx + \underline{\int}_{A\cup B} (f\mathbf{1}_B)(x)\,dx \leqslant \underline{\int}_{A\cup B} (f\mathbf{1}_A + f\mathbf{1}_B)(x)\,dx \\ &= \underline{\int}_{A\cup B} \big(f\mathbf{1}_{A\cup B} - (-f\mathbf{1}_{A\cap B})\big)(x)\,dx \\ &\leqslant \underline{\int}_{A\cup B} f\mathbf{1}_{A\cup B}(x)\,dx - \underline{\int}_{A\cup B} (-f\mathbf{1}_{A\cap B})(x)\,dx \\ &= \underline{\int}_{A\cup B} f(x)\,dx - \underline{\int}_{A\cap B} (-f)(x)\,dx\end{aligned}$$

which, with the help of Proposition 6.41 (e), further implies that

$$\underline{\int}_A f(x)\,dx + \underline{\int}_B f(x)\,dx \leqslant \underline{\int}_{A\cup B} f(x)\,dx\,.$$

The case of the upper integral can be proved in a similar fashion. □

Having established Proposition 6.41, it is easy to see the following theorem (except (c)). The proof is left as an exercise.

Theorem 6.44. *Let $A \subseteq \mathbb{R}^n$ be a bounded set, $c \in \mathbb{R}$, and $f, g : A \to \mathbb{R}$ be Riemann integrable functions. Then*

(a) *$f \pm g$ is Riemann integrable, and* $\displaystyle\int_A (f \pm g)(x)\,dx = \int_A f(x)\,dx \pm \int_A g(x)\,dx$.

(b) *cf is Riemann integrable, and* $\displaystyle\int_A (cf)(x)\,dx = c\int_A f(x)\,dx$.

(c) *$|f|$ is Riemann integrable, and* $\displaystyle\Big|\int_A f(x)\,dx\Big| \leqslant \int_A |f(x)|dx$.

(d) *If $f \leqslant g$, then* $\displaystyle\int_A f(x)\,dx \leqslant \int_A g(x)\,dx$.

(e) *If A has volume and $|f| \leqslant M$, then* $\displaystyle\Big|\int_A f(x)\,dx\Big| \leqslant M\nu(A)$.

Theorem 6.45. *Let $A \subseteq \mathbb{R}^n$ be bounded, and $f : A \to \mathbb{R}$ be a Riemann integrable function.*

1. *If A has measure zero, then $\displaystyle\int_A f(x)dx = 0$.*

2. *If $f(x) \geqslant 0$ for all $x \in A$, and $\displaystyle\int_A f(x)dx = 0$, then the set $\{x \in A \,|\, f(x) \neq 0\}$ has measure zero.*

Proof. 1. We show that $L(f,\mathcal{P}) \leqslant 0 \leqslant U(f,\mathcal{P})$ for all partitions $\mathcal{P}$ of A. Let $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ be a partition of A. By Corollary 6.25, $\Delta_k \cap A^\complement \neq \varnothing$ for $k = 1, \cdots, N$; thus we must have $\inf\limits_{x\in\Delta_k} \bar{f}(x) \leqslant 0$ and $\sup\limits_{x\in\Delta_k} \bar{f}(x) \geqslant 0$. As a consequence, if $\mathcal{P}$ is a partition of A,

$$L(f,\mathcal{P}) = \sum_{k=1}^{N} \inf_{x\in\Delta_k} \bar{f}(x)\nu(\Delta_k) \leqslant 0 \quad \text{and} \quad U(f,\mathcal{P}) = \sum_{k=1}^{N} \sup_{x\in\Delta_k} \bar{f}(x)\nu(\Delta_k) \geqslant 0;$$

thus $\displaystyle\underline{\int_A} f(x)dx \leqslant 0 \leqslant \overline{\int_A} f(x)dx$. Since f is integrable on A, $\displaystyle\int_A f(x)dx = 0$.

2. Let $A_k = \{x \in A \,|\, f(x) \geqslant \frac{1}{k}\}$. We claim that A_k has measure zero for all $k \in \mathbb{N}$.

Let $\varepsilon > 0$ be given. Since $\displaystyle\int_A f(x)dx = 0$, there exists a partition $\mathcal{P}$ of A such that $U(f,\mathcal{P}) < \dfrac{\varepsilon}{k}$. Let $C = \{\Delta \in \mathcal{P} \,|\, \Delta \cap A_k \neq \varnothing\}$. Then $A_k \subseteq \bigcup\limits_{\Delta\in C} \Delta$, and

$$\frac{1}{k}\sum_{\Delta\in C} \nu(C) \leqslant \sum_{\Delta\in C} \sup_{x\in\Delta} \bar{f}(x)\nu(\Delta) \leqslant \sum_{\Delta\in\mathcal{P}} \sup_{x\in\Delta} \bar{f}(x)\nu(\Delta) = U(f,\mathcal{P}) < \frac{\varepsilon}{k}$$

which implies that $\sum\limits_{\Delta\in C} \nu(\Delta) < \varepsilon$. Therefore, A_k has measure zero; thus Theorem 6.26 implies that $A = \bigcup\limits_{k=1}^{\infty} A_k$ also has measure zero. □

Remark 6.46. Combining Corollary 6.36 and Theorem 6.45, we conclude that the integral of a bounded function on a compact set of measure zero is zero.

Remark 6.47. Let $A = \mathbb{Q} \cap [0,1]$ and $f : A \to \mathbb{R}$ be the constant function $f \equiv 1$. We have shown in Example 6.33 that f is not Riemann integrable. We note that A has no volume since $\partial A = [0,1]$ which is not a set of measure zero. However, A has measure zero since it consists of countable number of points.

1. Since f is continuous on A, the condition that A has volume in Corollary 6.35 cannot be removed.

2. Since A has measure zero, the condition that f is Riemann integrable in Theorem 6.45 cannot be removed.

Definition 6.48. Let $A \subseteq \mathbb{R}^n$ be a set and $f : A \to \mathbb{R}$ be a function. For $B \subseteq A$, the ***restriction of*** f ***to*** B is the function $f\big|_B : A \to \mathbb{R}$ given by $f|_B = f\mathbf{1}_B$. In other words,

$$f\big|_B(x) = \begin{cases} f(x) & \text{if } x \in B, \\ 0 & \text{if } x \in A\backslash B. \end{cases}$$

The following two theorems are direct consequences of (a) of Proposition 6.41 and Corollary 6.43.

Theorem 6.49. *Let A, B be bounded subsets of $\mathbb{R}^n$, $B \subseteq A$, and $f : A \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable on B if and only if $f|_B$ is Riemann integrable on A. In either cases,*

$$\int_A f\big|_B(x)\,dx = \int_B f(x)\,dx.$$

Theorem 6.50. *Let A, B be bounded subsets of $\mathbb{R}^n$ be such that $A \cap B$ has volume zero, and $f : A \cup B \to \mathbb{R}$ be bounded such that $f\big|_A$ and $f\big|_B$ are all Riemann integrable on $A \cup B$. Then f is Riemann integrable on $A \cup B$, and*

$$\int_{A\cup B} f(x)\,dx = \int_A f(x)\,dx + \int_B f(x)\,dx.$$

6.4 The Fubini Theorem

If $f : [a, b] \to \mathbb{R}$ is continuous, the fundamental theorem of Calculus can be applied to computed the integral of f on $[a, b]$. In this section, we focus on how the integral of f on $A \subseteq \mathbb{R}^n$, where $n \geqslant 2$, can be computed if the integral exists.

Definition 6.51. Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be bounded sets, and $f : A \times B \to \mathbb{R}$ be a bounded function. For each fixed $x \in A$, the lower integral of the function $f(x, \cdot) : B \to \mathbb{R}$ is denoted by $\underline{\int_B} f(x, y)\,dy$, and the upper integral of $f(x, \cdot) : B \to \mathbb{R}$ is denoted by $\overline{\int_B} f(x, y)\,dy$. If the upper integral and the lower integral of $f(x, \cdot) : B \to \mathbb{R}$ are the same at $x \in A$, we simply write $\displaystyle\int_B f(x,y)\,dy$ for the integrals of $f(x,\cdot)$ on B. The integrals $\displaystyle\overline{\int_A} f(x,y)\,dx$, $\displaystyle\underline{\int_A} f(x,y)\,dx$ and $\displaystyle\int_A f(x,y)\,dx$ are defined similarly.

Theorem 6.52 (Fubini's Theorem). *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be bounded sets, and $f : A \times B \to \mathbb{R}$ be a bounded function. For $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, write $z = (x,y)$. Then*

$$\underline{\int_{A\times B}} f(z)\,dz \leqslant \underline{\int_A} \Big(\underline{\int_B} f(x,y)dy\Big)dx \leqslant \overline{\int_A} \Big(\overline{\int_B} f(x,y)dy\Big)dx \leqslant \overline{\int_{A\times B}} f(z)\,dz, \tag{6.4.1a}$$

$$\underline{\int_{A\times B}} f(z)\,dz \leqslant \underline{\int_B} \Big(\underline{\int_A} f(x,y)dx\Big)dy \leqslant \overline{\int_B} \Big(\overline{\int_A} f(x,y)dx\Big)dy \leqslant \overline{\int_{A\times B}} f(z)\,dz. \tag{6.4.1b}$$

In particular, if $f : A \times B \to \mathbb{R}$ is Riemann integrable, then

$$\begin{aligned}\int_{A\times B} f(z)\,dz &= \int_A \Big(\underline{\int_B} f(x,y)dy\Big)dx = \int_A \Big(\overline{\int_B} f(x,y)dy\Big)dx \\ &= \int_B \Big(\underline{\int_A} f(x,y)dx\Big)dy = \int_B \Big(\overline{\int_A} f(x,y)dx\Big)dy.\end{aligned}$$

Proof. It suffices to prove (6.4.1a). Let $\varepsilon > 0$ be given. Choose a partition $\mathcal{P}$ of $A \times B$ such that $L(f,\mathcal{P}) > \displaystyle\underline{\int_{A\times B}} f(z)\,dz - \varepsilon$. Since $\mathcal{P}$ is a partition of $A \times B$, there exist partition $\mathcal{P}_A$ of A and partition $\mathcal{P}_B$ of B such that $\mathcal{P} = \big\{\Delta = R \times S \,\big|\, R \in \mathcal{P}_A, S \in \mathcal{P}_B\big\}$. By Proposition 6.41 and Corollary 6.43, we find that

$$\begin{aligned}\underline{\int_A} \Big(\underline{\int_B} f(x,y)\,dy\Big)dx &= \underline{\int_{\bigcup_{R\in\mathcal{P}_A} R}} \mathbf{1}_A(x)\Big(\underline{\int_{\bigcup_{S\in\mathcal{P}_B} S}} f(x,y)\mathbf{1}_B(y)\,dy\Big)dx \\ &\geqslant \sum_{R\in\mathcal{P}_A} \underline{\int_R} \Big(\sum_{S\in\mathcal{P}_B} \underline{\int_S} \overline{f}^{A\times B}(x,y)\,dy\Big)dx \\ &\geqslant \sum_{R\in\mathcal{P}_A} \sum_{S\in\mathcal{P}_B} \underline{\int_R} \Big(\underline{\int_S} \overline{f}^{A\times B}(x,y)\,dy\Big)dx \\ &\geqslant \sum_{R\in\mathcal{P}_A, S\in\mathcal{P}_B} \inf_{(x,y)\in R\times S} \overline{f}^{A\times B}(x,y)\nu_m(S)\nu_n(R) \\ &= \sum_{\Delta\in\mathcal{P}} \inf_{(x,y)\in\Delta} \overline{f}^{A\times B}(x,y)\nu_{n+m}(\Delta) = L(f,\mathcal{P}) > \underline{\int_{A\times B}} f(z)dz - \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that

$$\underline{\int_{A\times B}} f(z)\,dz \leqslant \underline{\int_B} \Big(\underline{\int_A} f(x,y)dx\Big)dy.$$

Similarly, $\displaystyle\overline{\int_A} \Big(\overline{\int_B} f(x,y)dy\Big)dx \leqslant \overline{\int_{A\times B}} f(z)\,dz$; thus (6.4.1a) is concluded. $\square$

Corollary 6.53. *Let $S \subseteq \mathbb{R}^n$ be a bounded set with volume, $\varphi_1, \varphi_2 : S \to \mathbb{R}$ be continuous maps such that $\varphi_1(x) \leqslant \varphi_2(x)$ for all $x \in S$, $A = \{(x,y) \in \mathbb{R}^n \times \mathbb{R} \,|\, x \in S, \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\}$, and $f : A \to \mathbb{R}$ be continuous. Then f is Riemann integrable on A, and*

$$\int_A f(x,y)\, d(x,y) = \int_S \Big(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\, dy \Big) dx\,. \tag{6.4.2}$$

Proof. Since ∂A has measure zero, and f is continuous on A, Corollary 6.35 implies that f is Riemann integrable on A. Let $m = \min\limits_{x \in S} \varphi_1(x)$ and $M = \max\limits_{x \in S} \varphi_2(x)$. Then $A \subseteq S \times [m, M]$; thus Theorem 6.50 and the Fubini Theorem imply that

$$\begin{aligned} \int_A f(x,y)\, d(x,y) &= \int_{S \times [m,M]} \overline{f}^A(x,y)\, d(x,y) = \underline{\int_S} \Big(\underline{\int_m^M} \overline{f}^A(x,y)\, dy \Big) dx \\ &= \overline{\int_S} \Big(\overline{\int_m^M} \overline{f}^A(x,y)\, dy \Big) dx\,. \end{aligned}$$

Noting that $[m, M]$ has a boundary of volume zero in $\mathbb{R}$, and for each $x \in S$, $\overline{f}^A(x,\cdot)$ is continuous except perhaps at $y = \varphi_1(x)$ and $y = \varphi_2(x)$, Corollary 6.35 implies that $\overline{f}^A(x,\cdot)$ is Riemann integrable on $[m, M]$ for each $x \in S$. Therefore, $\underline{\int_m^M} \overline{f}^A(x,y)\, dy = \overline{\int_m^M} \overline{f}^A(x,y)\, dy$ which further implies that

$$\int_A f(x,y)\, d(x,y) = \int_S \Big(\int_m^M \overline{f}^A(x,y)\, dy \Big) dx\,. \tag{6.4.3}$$

For each fixed $x \in S$, let $A_x = \{y \in \mathbb{R} \,|\, \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\}$. Then $\overline{f}^A(x,y) = f(x,y)\mathbf{1}_{A_x}(y)$ for all $(x,y) \in S \times [m, M]$ or equivalently, $\overline{f}^A(x,\cdot) = f(x,\cdot)|_{A_x}$ for all $x \in S$; thus Proposition 6.41 (a) implies that

$$\int_m^M \overline{f}^A(x,y)\, dy = \int_{A_x} f(x,y)\, dy = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\, dy \qquad \forall\, x \in S\,. \tag{6.4.4}$$

Combining (6.4.3) and (6.4.4), we conclude (6.4.2). □

Corollary 6.54. 1. *Let $\varphi_1, \varphi_2 : [a, b] \to \mathbb{R}$ be continuous maps such that $\varphi_1(x) \leqslant \varphi_2(x)$ for all $x \in [a, b]$, $A = \{(x,y) \,|\, a \leqslant x \leqslant b, \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\}$, and $f : A \to \mathbb{R}$ be continuous. Then f is Riemann integrable on A, and*

$$\int_A f(x,y)\, d\mathbb{A} = \int_a^b \Big(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\, dy \Big) dx\,.$$

2. *Let* $\psi_1, \psi_2 : [c,d] \to \mathbb{R}$ *be continuous maps such that* $\psi_1(y) \leqslant \psi_2(y)$ *for all* $y \in [c,d]$, $A = \big\{(x,y) \,\big|\, c \leqslant y \leqslant d, \psi_1(y) \leqslant x \leqslant \psi_2(y)\big\}$, *and* $f : A \to \mathbb{R}$ *be continuous. Then* f *is Riemann integrable on* A, *and*

$$\int_A f(x,y)\, d\mathbb{A} = \int_c^d \Big(\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\, dx \Big) dy .$$

Remark 6.55. To simplify the notation, sometimes we use $\displaystyle\int_S \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\, dydx$ to denote the iterated integral the iterated integral $\displaystyle\int_S \Big(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\, dy \Big) dx$. Similar notation applies to the upper and the lower integrals. For example, we also have $\displaystyle\overline{\int_a^b} \underline{\int_c^d} f(x,y)\, dydx = \overline{\int_a^b} \Big(\underline{\int_c^d} f(x,y)\, dy \Big) dx$.

Remark 6.56. For each $x \in [a,b]$, define $\varphi(x) = \displaystyle\underline{\int_c^d} f(x,y)\, dy$ and $\psi(x) = \displaystyle\overline{\int_c^d} f(x,y)\, dy$. Then $\varphi(x) \leqslant \psi(x)$ for all $x \in [a,b]$, and the Fubini Theorem implies that

$$\int_a^b \big[\psi(x) - \varphi(x)\big]\, dx = 0 .$$

By Theorem 6.45, the set $\big\{x \in [a,b] \,\big|\, \psi(x) - \varphi(x) \neq 0\big\}$ has measure zero. In other words, except on a set of measure zero, $f(x,\cdot)$ is Riemann integrable on $[c,d]$ if f is Riemann integrable on $[a,b] \times [c,d]$. This property can be rephrased as that "$f(x,\cdot)$ is Riemann integrable on $[c,d]$ for **almost every** $x \in [a,b]$ if f is Riemann integrable on the rectangle $[a,b] \times [c,d]$". Similarly, $f(\cdot,y)$ is Riemann integrable for almost every $y \in [c,d]$ if f is Riemann integrable on $[a,b] \times [c,d]$.

Remark 6.57. The integrability of f does not guarantee that $f(x,\cdot)$ or $f(\cdot,y)$ is Riemann integrable. In fact, there exists a function $f : [0,1] \times [0,1] \to \mathbb{R}$ such that f is Riemann integrable, $f(\cdot,y)$ is Riemann integrable for each $y \in [0,1]$, but $f(x,\cdot)$ is not Riemann integrable for infinitely many $x \in [0,1]$. For example, let

$$f(x,y) = \begin{cases} 0 & \text{if } x = 0 \text{ or if } x \text{ or } y \text{ is irrational}, \\ \dfrac{1}{p} & \text{if } x, y \in \mathbb{Q} \text{ and } x = \dfrac{q}{p} \text{ with } (p,q) = 1 . \end{cases}$$

Then

1. For each $y \in [0,1]$, $f(\cdot, y)$ is continuous at all irrational numbers. Therefore, $f(\cdot, y)$ is Riemann integrable, and $\displaystyle\int_0^1 f(x,y)\,dx = \underline{\int_0^1} f(x,y)\,dx = 0$.

2. For $x = 0$ or $x \notin \mathbb{Q}$, $f(x,\cdot)$ is Riemann integrable, and $\displaystyle\int_0^1 f(x,y)\,dy = 0$.

3. If $x = \dfrac{q}{p}$ with $(p,q) = 1$, $f(x,\cdot)$ is nowhere continuous in $[0,1]$. In fact, for each $y_0 \in [0,1]$,
$$\lim_{\substack{y\to y_0\\ y\in\mathbb{Q}}} f(x,y) = \frac{1}{p} \quad \text{while} \quad \lim_{\substack{y\to y_0\\ y\notin\mathbb{Q}}} f(x,y) = 0;$$
thus the limit of $f(x,y)$ as $y \to y_0$ does not exist. Therefore, the Lebesgue theorem implies that $f(x,\cdot)$ is not Riemann integrable if $x \in \mathbb{Q} \cap (0,1]$. On the other hand, for $x = \dfrac{q}{p}$ with $(p,q) = 1$ we have
$$\underline{\int_0^1} f(x,y)\,dy = 0 \quad \text{and} \quad \overline{\int_0^1} f(x,y)\,dy = \frac{1}{p}.$$

4. Define $\varphi(x) = \displaystyle\underline{\int_0^1} f(x,y)\,dy$ and $\psi(x) = \displaystyle\overline{\int_0^1} f(x,y)\,dy$. Then 2 and 3 imply that φ and ψ are Riemann integrable on $[0,1]$, and $\displaystyle\int_0^1 \varphi(x)dx = \int_0^1 \psi(x)dx = 0$.

5. For each $a \notin \mathbb{Q} \cap [0,1]$ and $b \in [0,1]$, f is continuous at (a,b). In fact, for any given $\varepsilon > 0$, there exists a prime number p such that $\dfrac{1}{p} < \varepsilon$. Let
$$\delta = \min\Big\{ \big|a - \frac{\ell}{k}\big| \,\Big|\, 0 \leqslant \ell \leqslant k \leqslant p, k \in \mathbb{N}, \ell \in \mathbb{N} \cup \{0\} \Big\}.$$
Then $\delta > 0$, and if $(x,y) \in B\big((a,b),\delta\big) \cap ([0,1] \times [0,1])$, we have
$$\big|f(x,y) - f(a,b)\big| = \big|f(x,y)\big| < \frac{1}{p} < \varepsilon\,,$$
where we use the fact that if $(x,y) \in B\big((a,b),\delta\big)$ and $x \in \mathbb{Q}$, then $x = \dfrac{\ell}{k}$ (in reduced form) for some $k > p$.

As a consequence, $\big\{(a,b) \in \mathbb{R}^2 \,\big|\, \bar{f} \text{ is discontinuous at } (a,b)\big\} \subseteq \mathbb{Q} \times [0,1]$. Since $\mathbb{Q} \times [0,1]$ is a countable union of measure zero sets, it has measure zero; thus f is Riemann integrable by the Lebesgue theorem. The Fubini theorem then implies that
$$\int_{[0,1]\times[0,1]} f(x,y)\,d\mathbb{A} = \int_0^1 \int_0^1 f(x,y)\,dxdy = 0\,.$$

Remark 6.58. The integrability of $f(x,\cdot)$ and $f(\cdot,y)$ does not guarantee the integrability of f. In fact, there exists a bounded function $f:[0,1]\times[0,1]\to\mathbb{R}$ such that $f(x,\cdot)$ and $f(\cdot,y)$ are both Riemann integrable on $[0,1]$, but f is not Riemann integrable on $[0,1]\times[0,1]$. For example, let

$$f(x,y)=\begin{cases}1 & \text{if } (x,y)=\big(\frac{k}{2^n},\frac{\ell}{2^n}\big),\ 0<k,\ell<2^n \text{ odd numbers, } n\in\mathbb{N},\\ 0 & \text{otherwise.}\end{cases}$$

Then for each $x\in[0,1]$, $f(x,\cdot)$ only has finite number of discontinuities; thus $f(x,\cdot)$ is Riemann integrable, and

$$\int_0^1 f(x,y)\,dy=0.$$

Similarly, $f(\cdot,y)$ is Riemann integrable, and $\displaystyle\int_0^1 f(x,y)\,dx=0$. As a consequence,

$$\int_0^1\int_0^1 f(x,y)\,dydx=\int_0^1\int_0^1 f(x,y)\,dxdy=0.$$

However, note that f is nowhere continuous on $[0,1]\times[0,1]$; thus the Lebesgue theorem implies that f is not Riemann integrable. One can also see this by the fact that $U(f,\mathcal{P})=1$ and $L(f,\mathcal{P})=0$ for all partition of $[0,1]\times[0,1]$.

Example 6.59. Let $A=\big\{(x,y)\in\mathbb{R}^2\,\big|\,0\leqslant x\leqslant 1, x\leqslant y\leqslant 1\big\}$, and $f:A\to\mathbb{R}$ be given by $f(x,y)=xy$. Then Corollary 6.54 implies that

$$\int_A f(x,y)\,d\mathbb{A}=\int_0^1\Big(\int_x^1 xy\,dy\Big)dx=\int_0^1\frac{xy^2}{2}\Big|_{y=x}^{y=1}dx=\int_0^1\Big(\frac{x}{2}-\frac{x^3}{2}\Big)dx=\frac{1}{4}-\frac{1}{8}=\frac{1}{8}.$$

On the other hand, since $A=\big\{(x,y)\in\mathbb{R}^2\,\big|\,0\leqslant y\leqslant 1, 0\leqslant x\leqslant y\big\}$, we can also evaluate the integral of f on A by

$$\int_A xy\,d\mathbb{A}=\int_0^1\Big(\int_0^y xy\,dx\Big)dy=\int_0^1\frac{x^2y}{2}\Big|_{x=0}^{x=y}dy=\int_0^1\frac{y^3}{2}dy=\frac{1}{8}.$$

Example 6.60. Let $A=\big\{(x,y)\in\mathbb{R}^2\,\big|\,0\leqslant x\leqslant 1,\sqrt{x}\leqslant y\leqslant 1\big\}$, and $f:A\to\mathbb{R}$ be given by $f(x,y)=e^{y^3}$. Then Corollary 6.54 implies that

$$\int_A f(x,y)\,d\mathbb{A}=\int_0^1\Big(\int_{\sqrt{x}}^1 e^{y^3}dy\Big)dx.$$

Since we do not know how to compute the inner integral, we look for another way of finding the integral. Observing that $A = \{(x, y) \in \mathbb{R}^2 \,|\, 0 \leqslant y \leqslant 1, 0 \leqslant x \leqslant y^2\}$, we have

$$\int_A f(x,y)\, d\mathbb{A} = \int_0^1 \Big(\int_0^{y^2} e^{y^3}\, dx\Big) dy = \int_0^1 y^2 e^{y^3} dy = \frac{1}{3} e^{y^3}\Big|_{y=0}^{y=1} = \frac{e-1}{3}.$$

Example 6.61. Let $A \subseteq \mathbb{R}^3$ be the set $\{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0, \text{and } x_1 + x_2 + x_3 \leqslant 1\}$, and $f : A \to \mathbb{R}$ be given by $f(x_1, x_2, x_3) = (x_1 + x_2 + x_3)^2$. Let $S = [0,1] \times [0,1] \times [0,1]$, and $\bar{f} : \mathbb{R}^3 \to \mathbb{R}$ be the extension of f by zero outside A. Then Corollary 6.35 implies that f is Riemann integrable (since ∂A has measure zero). Write $\widehat{x}_1 = (x_2, x_3)$, $\widehat{x}_2 = (x_1, x_3)$ and $\widehat{x}_3 = (x_1, x_2)$. Theorem 6.49 implies that

$$\int_A f(x)dx = \int_S \bar{f}(x)dx\,,$$

and Theorem 6.52 implies that

$$\int_S \bar{f}(x)dx = \int_{[0,1]} \Big(\int_{[0,1]\times[0,1]} \bar{f}(\widehat{x}_3, x_3) d\widehat{x}_3\Big) dx_3\,.$$

Let $A_{x_3} = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_1 \geqslant 0, x_2 \geqslant 0, x_1 + x_2 \leqslant 1 - x_3\}$. Then for each $x_3 \in [0,1]$,

$$\int_{[0,1]\times[0,1]} \bar{f}(\widehat{x}_3, x_3) d\widehat{x}_3 = \int_{A_{x_3}} f(\widehat{x}_3, x_3) d\widehat{x}_3 = \int_0^{1-x_3} \Big(\int_0^{1-x_3-x_2} f(x_1, x_2, x_3) dx_1\Big) dx_2\,.$$

Computing the iterated integral, we find that

$$\begin{aligned}
\int_A f(x)dx &= \int_0^1 \Big[\int_0^{1-x_3} \Big(\int_0^{1-x_3-x_2} (x_1 + x_2 + x_3)^2 dx_1\Big) dx_2\Big] dx_3 \\
&= \int_0^1 \Big[\int_0^{1-x_3} \frac{(x_1 + x_2 + x_3)^3}{3}\Big|_{x_1=0}^{x_1=1-x_3-x_2} dx_2\Big] dx_3 \\
&= \int_0^1 \Big[\int_0^{1-x_3} \Big(\frac{1}{3} - \frac{(x_2 + x_3)^3}{3}\Big) dx_2\Big] dx_3 \\
&= \int_0^1 \Big(\frac{1}{4} - \frac{x_3}{3} + \frac{x_3^4}{12}\Big) dx_3 = \frac{1}{4} - \frac{1}{6} + \frac{1}{60} = \frac{15 - 10 + 1}{60} = \frac{1}{10}\,.
\end{aligned}$$

Example 6.62. In this example we compute the volume of the n-dimensional unit ball ω_n. By the Fubini theorem,

$$\omega_n = \int_{-1}^1 \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{n-1}^2}} dx_n \cdots dx_1\,.$$

Note that the integral $\displaystyle\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{n-1}^2}} dx_n \cdots dx_2$ is in fact $\omega_{n-1}(1-x_1^2)^{\frac{n-1}{2}}$; thus

$$\omega_n = \int_{-1}^{1} \omega_{n-1}(1-x^2)^{\frac{n-1}{2}}\, dx = 2\,\omega_{n-1} \int_0^{\frac{\pi}{2}} \cos^n \theta d\theta. \tag{6.4.5}$$

Integrating by parts,

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} \cos^n\theta\, d\theta &= \int_0^{\frac{\pi}{2}} \cos^{n-1}\theta\, d(\sin\theta) = \cos^{n-1}\theta \sin\theta\Big|_{\theta=0}^{\theta=\frac{\pi}{2}} + (n-1)\int_0^{\frac{\pi}{2}} \cos^{n-2}\theta \sin^2\theta\, d\theta \\
&= (n-1)\int_0^{\frac{\pi}{2}} \cos^{n-2}\theta(1-\cos^2\theta)\, d\theta
\end{aligned}$$

which implies that

$$\int_0^{\frac{\pi}{2}} \cos^n\theta\, d\theta = \frac{n-1}{n}\int_0^{\frac{\pi}{2}} \cos^{n-2}\theta\, d\theta.$$

As a consequence,

$$\int_0^{\frac{\pi}{2}} \cos^n\theta\, d\theta = \begin{cases} \dfrac{(n-1)(n-3)\cdots 2}{n(n-2)\cdots 3}\displaystyle\int_0^{\frac{\pi}{2}} \cos\theta\, d\theta & \text{if } n \text{ is odd}, \\[2ex] \dfrac{(n-1)(n-3)\cdots 1}{n(n-2)\cdots 2}\displaystyle\int_0^{\frac{\pi}{2}} d\theta & \text{if } n \text{ is even}; \end{cases}$$

thus the recursive formula (6.4.5) implies that $\omega_n = \dfrac{2\,\omega_{n-2}}{n}\pi$. Further computations shows that

$$\omega_n = \begin{cases} \dfrac{(2\pi)^{\frac{n-1}{2}}}{n(n-2)\cdots 3}\omega_1 & \text{if } n \text{ is odd}, \\[2ex] \dfrac{(2\pi)^{\frac{n-2}{2}}}{n(n-2)\cdots 4}\omega_2 & \text{if } n \text{ is even}. \end{cases}$$

Let Γ be the Gamma function defined by $\Gamma(t) = \displaystyle\int_0^\infty x^{t-1}e^{-x}\, dx$ for $t > 0$. Then $\Gamma(x+1) = x\Gamma(x)$ for all $x > 0$, $\Gamma(1) = 1$ and $\Gamma\big(\frac{1}{2}\big) = \sqrt{\pi}$. By the fact that $\omega_1 = 2$ and $\omega_2 = \pi$, we can express ω_n as

$$\omega_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\big(\frac{n+2}{2}\big)}.$$

6.5 The Monotone and Bounded Convergence Theorems

In the following, we introduce two very important theorems in the theory of integration of functions. Before proceeding, we establish the following two lemmas.

Lemma 6.63. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Then for each $\varepsilon > 0$, there exist continuous functions $g, h : [a,b] \to \mathbb{R}$ such that $\inf\limits_{x\in[a,b]} f(x) \leqslant g \leqslant f \leqslant h \leqslant \sup\limits_{x\in[a,b]} f(x)$ and*

$$\underline{\int_a^b} f(x)\,dx < \int_a^b g(x)\,dx + \varepsilon \qquad \textit{and} \qquad \overline{\int_a^b} f(x)\,dx > \int_a^b h(x)\,dx - \varepsilon\,.$$

Proof. We only prove the case of lower integral since the proof of the counter-part is similar.

Let $\varepsilon > 0$ be given, and $\mathcal{P} = \big\{a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\big\}$ be a partition of $[a,b]$ such that $L(f,\mathcal{P}) > \underline{\int_a^b} f(x)dx - \frac{\varepsilon}{2}$. Let $s(x)$ be the step function given by

$$s(x) = \sum_{k=1}^{n-1}\Big(\inf_{x\in[x_{k-1},x_k]} f(x)\Big)\mathbf{1}_{[x_k-1,x_k)}(x) + \Big(\inf_{x\in[x_{n-1},b]} f(x)\Big)\mathbf{1}_{[x_{n-1},b]}(x)$$

which is a linear combination of characteristic functions. Then $\inf\limits_{x\in[a,b]} f(x) \leqslant s(x) \leqslant f(x)$ for all $x \in [a,b]$ and

$$\underline{\int_a^b} f(x)\,dx < \int_a^b s(x)\,dx + \frac{\varepsilon}{2} \tag{6.5.1}$$

since the integral of s on $[a,b]$ is exactly the lower sum $L(f,\mathcal{P})$. On the other hand, for such a simple function s we can always find a continuous function $g : [a,b] \to \mathbb{R}$ (for example, g can be a trapezoidal function) such that $\inf\limits_{x\in[a,b]} f(x) \leqslant g(x) \leqslant s(x)$ for all $x \in [a,b]$ and

$$\int_a^b s(x)\,dx < \int_a^b g(x)\,dx + \frac{\varepsilon}{2}. \tag{6.5.2}$$

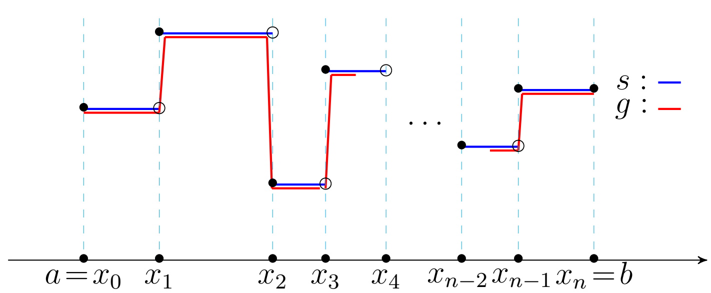


Figure 6.3: One way of constructing g given simple function s

The combination of (6.5.1) and (6.5.2) then concludes the lemma. □

Lemma 6.64. *Let $h_k : [a,b] \to \mathbb{R}$ be continuous for each $k \in \mathbb{N}$, and $h_k(x) \geqslant h_{k+1}(x)$ for each $k \in \mathbb{N}$ and $x \in [a,b]$. If $\lim\limits_{k\to\infty} h_k(x) = 0$ for each $x \in [a,b]$, then $\lim\limits_{k\to\infty} \int_a^b h_k(x)\,dx = 0$.*

Proof. For each $k \in \mathbb{N}$ define $c_k = \max\limits_{x\in[a,b]} h_k(x)$. Then $\{c_k\}_{k=1}^\infty$ is a decreasing sequence of non-negative real numbers, so $\lim\limits_{k\to\infty} c_k$ exists and is non-negative. Since $\displaystyle\int_a^b h_k(x)\,dx \leqslant c_k(b-a)$, it suffices to show that $\lim\limits_{k\to\infty} c_k = 0$.

Suppose the contrary that $\lim\limits_{k\to\infty} c_k = 2\delta$ for some $\delta > 0$. Then there exists $N > 0$ such that $c_k \geqslant \delta$ for all $k \geqslant N$. Define $F_k = \big\{x \in [a,b] \,\big|\, h_k(x) \geqslant \delta\big\}$. Then

1. F_k is closed for each $k \in \mathbb{N}$ by Theorem 4.14.

2. $F_k \supseteq F_{k+1}$ for each $k \in \mathbb{N}$.

3. $F_k \neq \varnothing$ for each $k \geqslant N$.

Therefore, by the nested set property we have $\bigcap\limits_{k=N}^\infty F_k \neq \varnothing$ $\Big($or otherwise $\bigcup\limits_{k=N}^\infty F_k^\complement$ is an open cover of compact set $[a,b]$ which, using the finite subcover property, implies that there exists $F_m \subseteq [a,b]^\complement$, a contradiction$\Big)$. This then implies that there exists $c \in [a,b]$ such that $h_k(c) \geqslant \delta$ for all $k \geqslant N$ which contradicts to the condition that $\lim\limits_{k\to\infty} h_k(x) = 0$ for all $x \in [a,b]$. Therefore, $\lim\limits_{k\to\infty} c_k = 0$. ▫

Theorem 6.65. *Let $\{f_k\}_{k=1}^\infty$ be a decreasing sequence of bounded functions on $[a,b]$; that is, for each $k \in \mathbb{N}$, $f_k(x) \geqslant f_{k+1}(x)$ for all $x \in [a,b]$ and f_k is bounded. If $\lim\limits_{k\to\infty} f_k(x) = 0$ for all $x \in [a,b]$, then*

$$\lim_{k\to\infty} \underline{\int_a^b} f_k(x)\,dx = 0 \,\Big(= \underline{\int_a^b} \lim_{k\to\infty} f_k(x)\,dx\Big).$$

Proof. Let $\varepsilon > 0$ be given. By Lemma 6.63, for each $k \in \mathbb{N}$ there exists a continuous function $g_k : [a,b] \to \mathbb{R}$ such that $0 \leqslant g_k \leqslant f_k$ and

$$\underline{\int_a^b} f_k(x)\,dx < \int_a^b g_k(x)\,dx + \frac{\varepsilon}{2^{k+1}}. \tag{6.5.3}$$

Define $h_k = \min\{g_1, \cdots, g_k\}$. Then h_k is continuous on $[a,b]$, $h_k \geqslant h_{k+1}$ (that is, $\{h_k\}_{k=1}^\infty$ is a decreasing sequence of funtions), $0 \leqslant h_k \leqslant g_k \leqslant f_k$ for all $k \in \mathbb{N}$, and $\lim\limits_{k\to\infty} h_k(x) = 0$ for all $x \in [a,b]$. Therefore, Lemma 6.64 implies that there exists $N > 0$ such that

$$\int_a^b h_k(x)\,dx < \frac{\varepsilon}{4} \qquad \forall\, k \geqslant N. \tag{6.5.4}$$

On the other hand, for $1 \leqslant \ell \leqslant k$, $\max\{g_\ell, \cdots, g_k\} \leqslant \max\{f_\ell, \cdots, f_k\} = f_\ell$; thus

$$\int_a^b \big(\max\{g_\ell, \cdots, g_k\}(x) - g_\ell(x)\big)\,dx \leqslant \underline{\int_a^b} f_\ell(x)\,dx - \int_a^b g_\ell(x)\,dx < \frac{\varepsilon}{2^{\ell+1}}.$$

Moreover, for each $1 \leqslant j \leqslant k$ and $x \in [a,b]$,

$$\begin{aligned} 0 \leqslant g_k(x) = g_j(x) + \big(g_k(x) - g_j(x)\big) &\leqslant g_j(x) + \big(\max\big\{g_j(x), \cdots, g_k(x)\big\} - g_j\big) \\ &\leqslant g_j(x) + \sum_{\ell=1}^{k-1} \big(\max\big\{g_\ell, \cdots, g_k\big\}(x) - g_\ell(x)\big), \end{aligned}$$

so minimizing the right-hand side over all $1 \leqslant j \leqslant k$ implies that

$$0 \leqslant g_k(x) \leqslant h_k(x) + \sum_{\ell=1}^{k-1} \big(\max\{g_\ell, \cdots, g_k\}(x) - g_\ell(x)\big) \qquad \forall\, x \in [a,b].$$

As a consequence,

$$0 \leqslant \int_a^b g_k(x)\, dx \leqslant \int_a^b h_k(x)\, dx + \sum_{\ell=1}^{k-1} \frac{\varepsilon}{2^{\ell+1}} \leqslant \int_a^b h_k(x)\, dx + \frac{\varepsilon}{2};$$

thus (6.5.3) and (6.5.4) imply that

$$0 \leqslant \underline{\int_a^b} f_k(x) dx < \varepsilon \qquad \forall\, k \geqslant N. \qquad \square$$

Example 6.66. Let $\{q_1, q_2, \cdots, q_n, \cdots\}$ denote the rational numbers in $[0,1]$, and $f_k : [0,1] \to \mathbb{R}$ be defined by

$$f_k(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q}^\complement \cup \{q_1, q_2, \cdots, q_k\}, \\ 1 & \text{otherwise}. \end{cases}$$

Then $f_k \geqslant f_{k+1}$ and $\lim\limits_{k\to\infty} f_k(x) = 0$ for all $x \in [0,1]$. Note that f_k is not Riemann integrable on $[0,1]$ but $\underline{\int_0^1} f_k(x)\, dx = 0$ for all $k \in \mathbb{N}$; thus

$$\lim_{k\to\infty} \underline{\int_0^1} f_k(x)\, dx = 0.$$

Note that $\overline{\int_0^1} f_k(x)\, dx = 1$ for all $k \in \mathbb{N}$.

Corollary 6.67 (Monotone Convergence Theorem)**.** *Let $f_k, f : [a,b] \to \mathbb{R}$ be Riemann integrable on $[a,b]$, and $\lim\limits_{k\to\infty} f_k(x) = f(x)$ for all $x \in [a,b]$. Suppose that $f_k \leqslant f_{k+1}$ for all $k \in \mathbb{N}$. Then*

$$\int_a^b f(x)\, dx = \lim_{k\to\infty} \int_a^b f_k(x)\, dx.$$

Proof. Let $g_k = f - f_k$. Then $\{g_k\}_{k=1}^\infty$ is a decreasing sequence of bounded functions on $[a,b]$ (since $0 \leqslant g_k \leqslant f - f_1$) and $\lim\limits_{k\to\infty} g_k(x) = 0$ for all $x \in [a,b]$. Therefore, the integrability of f_k and f, as well as Theorem 6.65, imply that

$$\begin{aligned} 0 &= \lim_{k\to\infty} \underline{\int_a^b} g_k(x)\,dx = \lim_{k\to\infty} \underline{\int_a^b} (f-f_k)(x)\,dx = \lim_{k\to\infty} \int_a^b (f-f_k)(x)dx \\ &= \int_a^b f(x)\,dx - \lim_{k\to\infty} \int_a^b f_k(x)\,dx\,. \end{aligned}$$ □

Corollary 6.68 (Arzelà's Bounded Convergence Theorem)**.** *Let $f_k, f : [a,b] \to \mathbb{R}$ be Riemann integrable on $[a,b]$, and $\lim\limits_{k\to\infty} f_k(x) = f(x)$ for all $x \in [a,b]$. Suppose that there exists a constant $M > 0$ such that $\big|f_k(x)\big| \leqslant M$ for all $x \in [a,b]$ and $k \in \mathbb{N}$. Then*

$$\int_a^b f(x)\,dx = \lim_{k\to\infty} \int_a^b f_k(x)\,dx\,.$$

Proof. Let $\varepsilon > 0$ be given. For each $k \in \mathbb{N}$, define $g_k(x) = \sup\limits_{\ell \geqslant k} \big|f_\ell(x) - f(x)\big|$. Then $\{g_k\}_{k=1}^\infty$ is a decreasing sequence of bounded functions on $[a,b]$ and $\lim\limits_{k\to\infty} g_k(x) = 0$ for all $x \in [a,b]$. Therefore, Theorem 6.65 implies that there exists $N > 0$ such that

$$\underline{\int_a^b} g_k(x)\,dx < \varepsilon \qquad \forall\, k \geqslant N\,.$$

Therefore, by observing that $0 \leqslant \big|f_k(x) - f(x)\big| \leqslant g_k(x)$ for all $k \in \mathbb{N}$, by the integrability of f_k and f we conclude that

$$\int_a^b \big|f_k(x) - f(x)\big|\,dx = \underline{\int_a^b} \big|f_k(x) - f(x)\big|\,dx \leqslant \underline{\int_a^b} g_k(x)\,dx < \varepsilon \qquad \forall\, k \geqslant N\,.$$ □

Theorem 6.69 (Monotone Convergence Theorem)**.** *Let $A = [a_1,b_1] \times \cdots \times [a_n,b_n]$ be a rectangle in $\mathbb{R}^n$, $f_k, f : A \to \mathbb{R}$ be Riemann integrable on A and $\lim\limits_{k\to\infty} f_k(x) = f(x)$ for all $x \in A$. Suppose that $\{f_k\}_{k=1}^\infty$ is a monotone sequence of functions; that is, $f_k \leqslant f_{k+1}$ or $f_k \geqslant f_{k+1}$ for all $k \in \mathbb{N}$. Then*

$$\lim_{k\to\infty} \int_A f_k(x)\,dx = \int_A f(x)\,dx\,.$$

Proof. W.L.O.G. we assume that $f_k \geqslant f_{k+1}$ for all $k \in \mathbb{N}$. We first prove the case $n = 2$ and write $A = [a,b] \times [c,d]$. Define $g_k(x) = \underline{\int_c^d} \big(f_k(x,y) - f(x,y)\big)dy$. Then $g_k \geqslant g_{k+1}$ for all $k \in \mathbb{N}$. Moreover, Theorem 6.65 implies that $\lim\limits_{k\to\infty} g_k(x) = 0$, and the Fubini theorem (Theorem 6.52) implies that g_k is Riemann integrable on $[a,b]$ for all $k \in \mathbb{N}$. Therefore, by the monotone convergence theorem for functions of one variable (Corollary 6.67) we find that

$$\begin{aligned}0 &= \lim_{k\to\infty} \int_a^b g_k(x)\,dx = \lim_{k\to\infty} \int_a^b \Big(\underline{\int_c^d} \big(f_k(x,y) - f(x,y)\big)dy\Big)dx \\ &= \lim_{k\to\infty} \int_A \big(f_k(x,y) - f(x,y)\big)\,d(x,y).\end{aligned}$$

Now suppose that the conclusion holds for the case $n = N$. Then for $n = N+1$, write $A = R \times [c,d]$ for some rectangle R in $\mathbb{R}^N$, and define g_k by

$$g_k(x_1, \cdots, x_N) = \underline{\int_c^d} \big(f_k(x_1, \cdots, x_{N+1}) - f(x_1, \cdots, x_{N+1})\big)dx_{N+1}.$$

Then Theorem 6.65 again implies that $\{g_k\}_{k=1}^\infty$ converges monotonically to 0 on R, and the Fubini theorem (Theorem 6.52) implies that g_k is Riemann integrable on R for all $k \in \mathbb{N}$. Write $x' = (x_1, \cdots, x_N)$. Then the validity of the monotone convergence theorem for N-tuple integrals implies that

$$\begin{aligned}0 &= \lim_{k\to\infty} \int_R g_k(x')\,dx' = \lim_{k\to\infty} \int_R \Big(\underline{\int_c^d} \big(f_k(x', x_{N+1}) - f(x', x_{N+1})\big)dx_{N+1}\Big)\,dx' \\ &= \lim_{k\to\infty} \int_A \big(f_k(x) - f(x)\big)\,dx.\end{aligned}$$

□

Theorem 6.70 (Bounded Convergence Theorem)**.** *Let* $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$ *be a rectangle in* $\mathbb{R}^n$*,* $f_k, f : A \to \mathbb{R}$ *be Riemann integrable on* A *and* $\lim\limits_{k\to\infty} f_k(x) = f(x)$ *for all* $x \in A$*. Suppose that there exists a constant* $M > 0$ *such that* $|f_k(x)| \leqslant M$ *for all* $x \in A$ *and* $k \in \mathbb{N}$*. Then*

$$\lim_{k\to\infty} \int_A f_k(x)\,dx = \int_A f(x)\,dx.$$

Proof. For $1 \leqslant j \leqslant n$, let R_j be the rectangle $[a_j, b_j] \times \cdots \times [a_n, b_n]$, and define $g_k^{(j)} : R_j \to \mathbb{R}$ iteratively as follows:

1. $g_k^{(1)}(x) = \sup\limits_{\ell \geqslant k} |f_\ell(x) - f(x)|$;

2. for $1 \leqslant j \leqslant n-1$, $g_k^{(j+1)}(x_{j+1}, \cdots, x_n) = \underline{\int_{a_j}^{b_j}} g_k^{(j)}(x_j, x_{j+1} \cdots, x_n)\,dx_j$.

Then for each $1 \leqslant j \leqslant n$, $\big\{g_k^{(j)}\big\}_{k=1}^\infty$ is a decreasing sequence of bounded functions; that is,

$$g_k^{(j)}(x_j, \cdots, x_n) \geqslant g_{k+1}^{(j)}(x_j, \cdots, x_n) \qquad \forall\, k \in \mathbb{N}, 1 \leqslant j \leqslant n, (x_j, \cdots, x_n) \in R_j\,.$$

Moreover, for each $1 \leqslant j \leqslant n$, $\lim\limits_{k\to\infty} g_k^{(j)}(x) = 0$ for all $x \in R_j$. To see this, we note that by the fact that $\lim\limits_{k\to\infty} f_k(x) = f(x)$ for all $x \in A$,

$$\lim_{k\to\infty} g_k^{(1)}(x) = \lim_{k\to\infty} \sup_{\ell \geqslant k} \big|f_\ell(x) - f(x)\big| = \limsup_{k\to\infty} \big|f_\ell(x) - f(x)\big| = 0\,.$$

Assume that for some $1 \leqslant j \leqslant n-1$ such that $\lim\limits_{k\to\infty} g_k^{(j)}(x) = 0$ for all $x \in R_j$. Then

$$\lim_{k\to\infty} g_k^{(j+1)}(x_{j+1}, \cdots, x_n) = \lim_{k\to\infty} \underline{\int_{a_j}^{b_j}} g_k^{(j)}(x_j, x_{j+1}, \cdots, x_n)\, dx_j = 0$$

for all $(x_{j+1}, \cdots, x_n) \in R_{j+1}$ which shows that $\lim\limits_{k\to\infty} g_k^{(j+1)}(x) = 0$ for all $x \in R_{j+1}$. By induction we conclude that $\lim\limits_{k\to\infty} g_k^{(j)}(x) = 0$ for all $x \in R_j$; thus Theorem 6.65 and the Fubini theorem (Theorem 6.52) imply that

$$\begin{aligned}
0 &= \lim_{k\to\infty} \underline{\int_{a_n}^{b_n}} g_k^{(n)}(x_n)\, dx_n = \lim_{k\to\infty} \underline{\int_{a_n}^{b_n}} \cdots \underline{\int_{a_1}^{b_1}} \sup_{\ell \geqslant k} \big|f_\ell(x_1, \cdots, x_n) - f(x_1, \cdots, x_n)\big|\, dx_1 \cdots dx_n \\
&\geqslant \limsup_{k\to\infty} \underline{\int_{a_n}^{b_n}} \cdots \underline{\int_{a_1}^{b_1}} \big|f_k(x_1, \cdots, x_n) - f(x_1, \cdots, x_n)\big|\, dx_1 \cdots dx_n \\
&\geqslant \limsup_{k\to\infty} \underline{\int_A} \big|f_k(x) - f(x)\big| dx = \limsup_{k\to\infty} \int_A \big|f_k(x) - f(x)\big| dx\,.
\end{aligned}$$

□

Remark 6.71. 1. If A is a bounded set with volume, we can choose a rectangle $S \supseteq A$ and consider $g_k = \overline{f}_k^A$ as well as $g = \overline{f}^A$. Then $g_k, g : S \to \mathbb{R}$ satisfy the assumptions in Theorem 6.69 and 6.70; thus Proposition 6.41 implies that

$$\lim_{k\to\infty} \int_A f_k(x)\, dx = \lim_{k\to\infty} \int_R g_k(x)\, dx = \int_R g(x)\, dx = \int_A f(x)\, dx\,.$$

In other words, the Monotone Convergence Theorem and the Bounded Convergence Theorem also hold for more general domain A, or to be more precise, for bounded set A with volume.

2. The Monotone Convergence Theorem (MCT) can be viewed as a corollary of the Bounded Convergence Theorem (BCT) since under the assumptions of MCT, we can apply BCT $\big($choose $M = \max\big\{\sup\limits_{x\in A} f(x), \sup\limits_{x\in A} f_1(x)\big\}\big)$ directly to conclude the MCT. Here we prove MCT without the help of BCT to demonstrate the power of the Fubini Theorem.

6.6 Improper Integrals

The Riemann integral deals with the "integrals" of bounded functions on bounded sets; however, often times we need to integrate unbounded functions on unbounded sets, such as finding the area under an unbounded function above x-axis. The improper integral is an answer to this particular situation. We first consider improper integrals of non-negative functions. Let $A \subseteq \mathbb{R}^n$ be a set and $f : A \to \mathbb{R}$ be a non-negative function. If f is bounded but A is unbounded, to define the integral of f on A, it is natural to consider the limit

$$\lim_{k\to\infty} \int_{A\cap B(0,k)} f(x)\, dx\,.$$

We note that for this limit to make sense, it is required that the integral $\displaystyle\int_{A\cap B(0,k)} f(x)\, dx$ exists for all $k \in \mathbb{N}$. On the other hand, if A is bounded and f is unbounded, to define the integral of f on A it is also natural to consider the limit

$$\lim_{k\to\infty} \int_{A} (f \wedge k)(x)\, dx\,, \tag{6.6.1}$$

where $(f \wedge k)(x) = \min\{f(x), k\}$. Again, for the limit above to make sense, it is required that the integral $\displaystyle\int_A (f \wedge k)(x)\, dx$ exists for all $k \in \mathbb{N}$. In both cases, we look for generic conditions (independent of k) that f and A have to satisfy so that

$$\int_{A\cap B(0,k)} f(x)\, dx \quad \text{and} \quad \int_A (f \wedge k)(x)\, dx$$

are well-defined for all $k \in \mathbb{N}$, and we have the following

Definition 6.72 (Riemann measurable sets and functions)**.**

1. A set $A \subseteq \mathbb{R}^n$ is said to be ***Riemann measurable*** if ∂A has measure zero.
2. A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ is said to be ***Riemann measurable*** if

$$\text{the set } \big\{x \in A \,\big|\, f \text{ is discontinuous at } x\big\} \text{ has measure zero in } \mathbb{R}^n.$$

Adopting this definition, the Lebesgue theorem shows that if A is bounded and Riemann measurable, then $f : A \to \mathbb{R}$ is Riemann integrable on A if and only if f is bounded and Riemann measurable. Since the improper integrals deal with integrals of possibly unbounded functions on possibly unbounded sets, in view of the Lebesgue theorem it is quite natural to consider removing the boundedness but keeping the Riemann measurability of A and f in order to define the improper integrals. Observe that

1. If A is Riemann measure, then $A \cap B(0,k)$ is Riemann measurable for all $k \in \mathbb{N}$ since
$$\partial(A \cap B(0,k)) \subseteq \partial A \cup \partial B(0,k).$$

2. If f is Riemann measure, then $f \wedge k$ is Riemann measurable for all $k \in \mathbb{N}$. In fact, since the function $F : \mathbb{R}^2 \to \mathbb{R}$ defined by $F(x,y) = \max\{x,y\}$ is continuous, if f is continuous at x, then for all $k \in \mathbb{N}$,
$$\lim_{y\to x}(f \wedge k)(y) = \lim_{y\to x} F(f(y),k) = F(f(x),k) = (f \wedge k)(x).$$
Therefore, for all $k \in \mathbb{N}$,
$$\{x \in A \,|\, f \wedge k \text{ is discontinuous at } x\} \subseteq \{x \in A \,|\, f \text{ is discontinuous at } x\}.$$

In other words, if $A \subseteq \mathbb{R}^n$ is a Riemann measurable set and $f : A \to \mathbb{R}$ is a Riemann measurable function.

1. If f is bounded, then $\displaystyle\int_{A\cap B(0,k)} f(x)\,dx$ exists for all $k \in \mathbb{N}$.

2. If A is bounded, then $\displaystyle\int_A (f \wedge k)(x)\,dx$ exists for all $k \in \mathbb{N}$.

How about the case that A is an unbounded set and f is an unbounded function? From the discussion above, it is natural to consider the limit of $\displaystyle\int_{A\cap B(0,k)} (f \wedge k)(x)\,dx$ as $k \to \infty$. We note that if $A \subseteq \mathbb{R}^n$ is a Riemann measurable set and $f : A \to \mathbb{R}$ is a Riemann measurable function, then $\displaystyle\int_{A\cap B(0,k)} (f \wedge k)(x)\,dx$ exists for all $k \in \mathbb{N}$. This motivates the following

Definition 6.73. Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f : A \to \mathbb{R}$ be a non-negative Riemann measurable function. f is said to be ***integrable*** on A if the limit
$$\int_A f(x)\,dx \equiv \lim_{k\to\infty} \int_{A\cap B(0,k)} (f \wedge k)(x)\,dx \tag{6.6.2}$$
is finite, and in such a case $\displaystyle\int_A f(x)\,dx$ is called the integral of f on A.

Remark 6.74. 1. For non-negative function $f : A \to \mathbb{R}$ (with f and A satisfying assumptions in Definition 6.73), if the limit $\displaystyle\int_{A\cap B(0,k)} f(x)\,dx$ is infinite, we still call $\displaystyle\int_A f(x)\,dx$ the integral of f on A. However, in this case f is not integrable on A.

2. Let $A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}$ be given in Definition 6.73. If $F \subseteq A$ is a Riemann measurable set with measure zero, then
$$\int_F f(x)\,dx = \lim_{k\to\infty} \int_{F\cap B(0,k)} (f \wedge k)(x)\,dx = 0\,,$$
where Theorem 6.45 is used to evaluate the integral.

3. By the Monotone Convergence Theorem (Theorem 6.69), (6.6.2) always holds if $f : A \to \mathbb{R}$ is Riemann integrable; thus Riemann integrable functions are integrable. **From now on, when talking about integrability, it could refer to Riemann integrable functions as well.**

Remark 6.75. When $f : A \to \mathbb{R}$ is unbounded, instead of (6.6.1) one might want to define the improper integral of f on A as
$$\lim_{k\to\infty} \int_A f_k(x)\,dx\,,$$
where
$$f_k(x) = \big(f\mathbf{1}_{\{f\leqslant k\}}\big)(x) = \begin{cases} f(x) & \text{if } f(x) \leqslant k\,, \\ 0 & \text{otherwise}. \end{cases}$$
The sequence $\{f_k\}_{k=1}^\infty$ still monotonically converges to f; however, it is not easy to see if the collection of points of discontinuity of f_k has measure zero since the set $\{x \in A \,|\, f(x) = k\}$ could be large. In other words, by defining f_k in this way we do not know the integrability of f_k on A; thus it is meaningless to define the improper integral as the limit of $\displaystyle\int_A f_k(x)\,dx$.

Example 6.76. Let $f : [1,\infty) \to \mathbb{R}$ be given by $f(x) = x^p$ for some $p \in \mathbb{R}$. If $p > 0$, then f is unbounded, and in this case
$$(f \wedge k)(x) = \begin{cases} x^p & \text{if } 1 \leqslant x \leqslant k^{\frac{1}{p}}\,, \\ k & \text{if } x > k^{\frac{1}{p}}\,; \end{cases}$$
thus for $p \geqslant 1$
$$\int_{[1,\infty)\cap(-k,k)} (f\wedge k)(x)\,dx = \int_1^{k^{\frac{1}{p}}} x^p dx + \int_{k^{\frac{1}{p}}}^k k dx = \frac{1}{p+1}(k^{1+\frac{1}{p}} - 1) + k(k - k^{\frac{1}{p}})$$
and for $0 < p < 1$,
$$\int_{[1,\infty)\cap(-k,k)} (f\wedge k)(x)\,dx = \int_1^k x^p dx = \frac{1}{p+1}(k^{p+1} - 1)\,.$$

In both cases, the limit (as $k \to \infty$) do not exist.

When $p \leqslant 0$, f is bounded by 1 on $[1,\infty)$. Therefore,

$$\int_{[1,\infty)\cap(-k,k)} (f \wedge k)(x)\,dx = \int_1^k f(x)dx = \begin{cases} \dfrac{1}{p+1}(k^{p+1}-1) & \text{if } p \neq -1\,, \\ \log k & \text{if } p = -1\,. \end{cases}$$

It is easy to see that the limit (as $k \to \infty$) exists only when $p < -1$. Therefore, f is integrable on $(0,1)$ if and only if $p < -1$, and in this case

$$\int_{[1,\infty)} f(x)\,dx = \lim_{k\to\infty} \frac{1}{p+1}(k^{p+1}-1) = -\frac{1}{p+1}\,.$$

Example 6.77. Let $f : (0,1) \to \mathbb{R}$ be given by $f(x) = x^p$ for some $p \in \mathbb{R}$. If $p \geqslant 0$, f is continuous on $(0,1)$, so f is Riemann integrable on $(0,1)$. If $p < 0$, f is unbounded on $(0,1)$, so the Riemann integral of f no longer makes sense. Nevertheless, we can find the improper integral of f using (6.6.2): for each $k \in \mathbb{N}$,

$$(f \wedge k)(x) = \begin{cases} x^p & \text{if } x \geqslant k^{\frac{1}{p}}\,, \\ k & \text{if } 0 < x < k^{\frac{1}{p}}\,; \end{cases}$$

thus

$$\int_0^1 (f \wedge k)(x)\,dx = \int_0^{k^{\frac{1}{p}}} k dx + \int_{k^{\frac{1}{p}}}^1 x^p dx = \begin{cases} \dfrac{1}{p+1}(pk^{1+\frac{1}{p}}+1) & \text{if } p \neq -1\,, \\ 1+\log k & \text{if } p = -1\,. \end{cases}$$

It is easy to see that the limit (as $k \to \infty$) exists only when $p > -1$. Therefore, f is integrable on $(0,1)$ if $p > -1$, and in this case

$$\int_{(0,1]} f(x)dx = \lim_{k\to\infty} \frac{1}{p+1}(pk^{1+\frac{1}{p}}+1) = \frac{1}{p+1}\,.$$

The following three propositions are generalization of corresponding results in Riemann integrals.

Proposition 6.78.

1. *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f, g : A \to \mathbb{R}$ be non-negative, Riemann measurable functions. If $f \leqslant g$, then*

$$\int_A f(x)\,dx \leqslant \int_A g(x)\,dx\,.$$

2. *Let $A, B \subseteq \mathbb{R}^n$ be Riemann measurable sets, and $f : A \cup B \to \mathbb{R}$ be a non-negative, Riemann measurable function. If $A \subseteq B$, then*
$$\int_A f(x)\,dx \leqslant \int_B f(x)\,dx\,.$$

Proof. The first case follows from that $\displaystyle\int_{A\cap B(0,k)} (f \wedge k)(x)\,dx \leqslant \int_{A\cap B(0,k)} (g\wedge k)(x)\,dx$, while the second case follows from that $\displaystyle\int_{A\cap B(0,k)} (f\wedge k)(x)\,dx \leqslant \int_{B\cap B(0,k)} (f\wedge k)(x)\,dx$. □

Corollary 6.79. *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f : A \to \mathbb{R}$ be a non-negative Riemann measurable function. Then*
$$\int_A f(x)\,dx = \lim_{k\to\infty} \int_{A\cap B(0,k)} f(x)\,dx = \lim_{k\to\infty} \int_A (f\wedge k)(x)\,dx\,. \tag{6.6.3}$$

Proof. For each $k \in \mathbb{N}$, Proposition 6.78 implies that
$$\int_{A\cap B(0,k)} (f\wedge k)(x)\,dx \leqslant \int_{A\cap B(0,k)} f(x)\,dx \leqslant \int_A f(x)\,dx\,,$$
$$\int_{A\cap B(0,k)} (f\wedge k)(x)\,dx \leqslant \int_A (f\wedge k)(x)\,dx \leqslant \int_A f(x)\,dx\,.$$
The conclusion follows from passing to the limit as $k \to \infty$. □

Corollary 6.80. *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f : A \to \mathbb{R}$ be a non-negative Riemann measurable function. Then for all $\alpha > 0$,*
$$\int_A (\alpha f)(x)\,dx = \alpha \int_A f(x)\,dx\,.$$

Proof. By Corollary 6.79,
$$\begin{aligned}\int_A (\alpha f)(x)\,dx &= \lim_{k\to\infty} \int_{A\cap B(0,\alpha k)} \big[(\alpha f)\wedge(\alpha k)\big](x)\,dx = \lim_{k\to\infty} \int_{A\cap B(0,\alpha k)} \alpha(f\wedge k)(x)\,dx \\ &= \alpha \lim_{k\to\infty} \int_{A\cap B(0,\alpha k)} (f\wedge k)(x)\,dx = \alpha \int_A f(x)\,dx\,.\end{aligned}$$
□

Proposition 6.81. *Let $A, B \subseteq \mathbb{R}^n$ be Riemann measurable sets, and $f : A \cup B \to \mathbb{R}$ be a non-negative, Riemann measurable function. If $A \cap B$ has measure zero, then*
$$\int_{A\cup B} f(x)\,dx = \int_A f(x)\,dx + \int_B f(x)\,dx\,.$$

Proof. To simplify the notation, for each $k \in \mathbb{N}$ we let $f_k = f \wedge k$, and $A_k = A \cap B(0,k)$ as well as $B_k = B \cap B(0,k)$. Then

$$\begin{aligned}
\partial(A_k \cup B_k) &= \partial\big((A \cup B) \cap B(0,k)\big) \subseteq \partial(A \cup B) \cup \partial B(0,k) \subseteq \partial A \cup \partial B \cup \partial B(0,k)\,,\\
\partial(A_k \cap B_k) &= \partial\big((A \cap B) \cap B(0,k)\big) \subseteq \partial(A \cap B) \cup \partial B(0,k) \subseteq \partial A \cup \partial B \cup \partial B(0,k)\,;
\end{aligned}$$

thus under the assumptions of this proposition, $A_k \cup B_k$ and $A_k \cap B_k$ have volume for each $k \in \mathbb{N}$. Therefore, Corollary 6.35 implies that for each $k \in \mathbb{N}$, $f_k 1_{A_k}$, $f_k 1_{B_k}$ and $f_k 1_{A_k \cap B_k}$ are all Riemann integrable on $A_k \cup B_k$. Since $A_k \cap B_k$ has measure zero, Theorem 6.45 and 6.50 imply that

$$\begin{aligned}
\int_{(A\cup B)\cap B(0,k)} (f \wedge k)\, dx &= \int_{A_k \cup B_k} f_k(x)\, dx = \int_{A_k} f_k(x)\, dx + \int_{B_k} f_k(x)\, dx\\
&= \int_{A \cap B(0,k)} (f \wedge k)\, dx + \int_{B \cap B(0,k)} (f \wedge k)\, dx\,,
\end{aligned}$$

and the theorem is concluded by passing to the limit as $k \to \infty$. □

Proposition 6.82. *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f, g : A \to \mathbb{R}$ be non-negative, Riemann measurable functions. Then*

$$\int_A (f+g)(x)\, dx = \int_A f(x)\, dx + \int_A g(x)\, dx\,.$$

Proof. Note that if f, g are non-negative functions, then for all $k \in \mathbb{N}$,

$$\big[(f+g) \wedge k\big](x) \leqslant (f \wedge k)(x) + (g \wedge k)(x) \leqslant \big[(f+g) \wedge (2k)\big](x) \qquad \forall\, x \in A\,.$$

Therefore, Proposition 6.78 implies that for all $k \in \mathbb{N}$,

$$\begin{aligned}
\int_{A\cap B(0,k)} \big[(f+g) \wedge k\big](x)\, dx &\leqslant \int_{A\cap B(0,k)} \big[(f \wedge k)(x) + (g \wedge k)(x)\big]\, dx\\
\leqslant \int_{A\cap B(0,2k)} &\big[(f \wedge k)(x) + (g \wedge k)(x)\big]\, dx \leqslant \int_{B(0,2k)} \big[(f+g) \wedge (2k)\big](x)\, dx\,.
\end{aligned}$$

By Theorem 6.44, we obtain that

$$\begin{aligned}
\int_{A\cap B(0,k)} \big[(f+g) \wedge k\big](x)\, dx &\leqslant \int_{A\cap B(0,k)} (f \wedge k)(x)\, dx + \int_{A\cap B(0,k)} (g \wedge k)(x)\, dx\\
&\leqslant \int_{B(0,2k)} \big[(f+g) \wedge (2k)\big](x)\, dx\,,
\end{aligned}$$

and the conclusion follows from passing to the limit as $k \to \infty$. □

Those who are familiar with the improper integrals introduced in Calculus might be confused with the way we compute the improper integrals in Example 6.76 and 6.77. In fact, there are other ways of evaluating the improper integrals for functions of one variable, and the following theorem is useful for this particular purpose.

Theorem 6.83. 1. *Let $f : [a,\infty) \to \mathbb{R}$ be bounded, non-negative, and continuous except perhaps on a set of measure zero. Then*

$$\int_{[a,\infty)} f(x)\,dx = \lim_{R\to\infty}\int_a^R f(x)\,dx\,. \tag{6.6.4}$$

2. *Let $a \in \mathbb{R}$, $f : (a,b] \to \mathbb{R}$ be non-negative, bounded on $[a+\varepsilon, b]$ for all $\varepsilon > 0$, and continuous except perhaps on a set of measure zero. Then*

$$\int_{(a,b]} f(x)\,dx = \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b f(x)\,dx\,. \tag{6.6.5}$$

Proof. 1. Note that for $k \geqslant \max\big\{|a|, \sup_{x\in[a,\infty)} f(x)\big\}$,

$$\int_{[a,\infty)\cap(-k,k)} (f\wedge k)(x)\,dx = \int_a^k f(x)\,dx\,;$$

thus (6.6.4) is obtained by passing to the limit as $k \to \infty$.

2. For each $\varepsilon > 0$ sufficiently small,

$$\int_{a+\varepsilon}^b f(x)\,dx = \int_{(a,b]} (f\mathbf{1}_{[a+\varepsilon,b]})(x)\,dx \leqslant \int_{(a,b]} f(x)\,dx\,;$$

thus passing to the limit as $\varepsilon \to 0^+$, we find that

$$\lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b f(x)\,dx \leqslant \int_{(a,b]} f(x)\,dx\,. \tag{6.6.6}$$

On the other hand, note that the Monotone Convergence Theorem for Riemann integrable sequence of functions (Theorem 6.69) implies that

$$\int_a^b (f\wedge k)(x)\,dx = \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b (f\wedge k)(x)\,dx \leqslant \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b f(x)\,dx\,.$$

Passing to the limit as $k \to \infty$, we find that

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0^+}\int_{(a,b]\cap(-k,k)} (f\wedge k)(x)\,dx \leqslant \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b f(x)\,dx\,. \tag{6.6.7}$$

Combining (6.6.6) and (6.6.7), we concluded (6.6.5). ◻

Corollary 6.84. *Let $a \in \mathbb{R}$, $f : (a,\infty) \to \mathbb{R}$ be non-negative, bounded on $[a+\varepsilon,\infty)$ for all $\varepsilon > 0$, and continuous except perhaps on a set of measure zero. Then*

$$\int_{(a,\infty)} f(x)\,dx = \lim_{\substack{R\to\infty \\ \varepsilon\to 0^+}} \int_{a+\varepsilon}^{R} f(x)\,dx\,. \tag{6.6.8}$$

Proof. Let $a < b < \infty$. Then Theorem 6.81 implies that

$$\int_{(a,\infty)} f(x)\,dx = \int_{(a,b]} f(x)\,dx + \int_{[b,\infty)} f(x)\,dx\,;$$

thus we conclude from Theorem 6.83 that

$$\begin{aligned}
\int_{(a,\infty)} f(x)\,dx &= \lim_{\varepsilon\to 0^+} \int_{[a+\varepsilon,b]} f(x)\,dx + \lim_{R\to\infty} \int_{[b,R]} f(x)\,dx \\
&= \lim_{\substack{R\to\infty \\ \varepsilon\to 0^+}} \int_{[a+\varepsilon,b]} f(x)\,dx + \lim_{\substack{R\to\infty \\ \varepsilon\to 0^+}} \int_{[b,R]} f(x)\,dx \\
&= \lim_{\substack{R\to\infty \\ \varepsilon\to 0^+}} \Big(\int_{[a+\varepsilon,b]} f(x)\,dx + \int_{[b,R]} f(x)\,dx \Big) = \lim_{\substack{R\to\infty \\ \varepsilon\to 0^+}} \int_{a+\varepsilon}^{R} f(x)\,dx\,,
\end{aligned}$$

where the sum of the limits of two integrals is the same as the limit of sums of integrals since both integrals are increasing as $R \to \infty$ and $\varepsilon \to 0^+$. ◻

Remark 6.85. In view of (6.6.4) and (6.6.5), we also have the following notation for improper integrals for functions of one variable:

$$\int_a^\infty f(x)\,dx \equiv \int_{[a,\infty)} f(x)\,dx \qquad \text{and} \qquad \int_a^b f(x)\,dx = \int_{(a,b]} f(x)\,dx\,.$$

Example 6.86. Let $f(x) = x^p$ as in Example 6.76 and 6.77. Since

$$\int_1^R x^p\,dx = \begin{cases} \dfrac{1}{p+1}(R^{p+1} - 1) & \text{if } p \neq -1\,, \\ \log R & \text{if } p = -1\,, \end{cases}$$

and

$$\int_\varepsilon^1 x^p\,dx = \begin{cases} \dfrac{1}{p+1}(1 - \varepsilon^{1+p}) & \text{if } p \neq -1\,, \\ -\log\varepsilon & \text{if } p = -1\,, \end{cases}$$

by Theorem 6.83 we find that f is integrable on $[1,\infty)$ if and only if $p < -1$ and f is integrable on $(0,1]$ if and only if $p > -1$. These are the conclusions that we have obtained in Example 6.76 and 6.77.

Example 6.87 (The Gamma function)**.** For each $t > 0$, define $\Gamma(t) = \displaystyle\int_0^\infty x^{t-1}e^{-x}\,dx$.

1. For $1 \leqslant t < \infty$, the integrand is bounded and non-negative. In fact, $x^{t-1}e^{-x} \leqslant M_t e^{-\frac{x}{2}}$ for some constant $M_t > 0$ (we can choose $M_t = \sup\limits_{x\in[0,\infty)} x^{t-1}e^{-\frac{x}{2}}$). Since
$$\int_0^R x^{t-1}e^{-x}\,dx \leqslant \int_0^R M_t e^{-\frac{x}{2}}dx \leqslant -2M_t e^{-\frac{x}{2}}\Big|_{x=0}^{x=R} \leqslant 2M_t < \infty\,;$$
we find that $\Gamma(t)$ is well-defined for $1 \leqslant t < \infty$.

2. For $0 < t < 1$, the integrand is unbounded near 0; thus by Theorem 6.81 we rewrite
$$\int_0^\infty x^{t-1}e^{-x}\,dx = \int_0^1 x^{t-1}e^{-x}\,dx + \int_1^\infty x^{t-1}e^{-\frac{x}{2}}e^{-\frac{x}{2}}dx\,.$$
Since $x^{t-1}e^{-x} \leqslant x^{t-1}$ on $(0,1]$ and $x^{t-1}e^{-x} \leqslant e^{-x}$ on $[1,\infty)$, for $\varepsilon > 0$ we have
$$\int_\varepsilon^1 x^{t-1}e^{-x}dx \leqslant \int_\varepsilon^1 x^{t-1}dx = \frac{1}{t}x^t\Big|_{x=\varepsilon}^1 = \frac{1}{t}(1-\varepsilon^t) \leqslant \frac{1}{t}$$
and for $R > 1$,
$$\int_1^R x^{t-1}e^{-x}dx \leqslant \int_1^R e^{-x}dx = -e^{-x}\Big|_{x=1}^{x=R} = e^{-1} - e^{-R} \leqslant e^{-1}\,.$$
Therefore, $\Gamma(t)$ is also well-defined for $0 < t < 1$.

The following theorem provides different ways of computing the improper (multiple) integrals.

Theorem 6.88. *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f : A \to \mathbb{R}$ be a non-negative, Riemann measurable function. Then f is integrable on A if and only if for each sequence $\{B_k\}_{k=1}^\infty \subseteq \mathbb{R}^n$ of bounded sets with volume satisfying*

1. *$B_k \subseteq B_{k+1}$ for all $k \in \mathbb{N}$;*

2. *for all $R > 0$ we have $B(0,R) \subseteq B_k$ for sufficient large $k \in \mathbb{N}$;*

the limit $\lim\limits_{k\to\infty} \displaystyle\int_{A\cap B_k} (f \wedge k)(x)\,dx$ exists.

Proof. "$\Leftarrow$" Simply choose $B_k = B(0,k)$ to conclude the integrability of f on A.

"$\Rightarrow$" For each $\ell \in \mathbb{N}$, there exists $N(\ell) \geqslant \ell$ such that $B(0,\ell) \subseteq B_k$ for all $k \geqslant N(\ell)$. Then

$$\int_{A\cap B(0,\ell)} (f \wedge \ell)(x)\, dx \leqslant \int_{A\cap B_k} (f \wedge \ell)(x)\, dx \leqslant \int_{A\cap B_k} (f \wedge k)(x)\, dx \qquad \forall\, k \geqslant N(\ell)\,.$$

Since $\displaystyle\int_{A\cap B_k} (f \wedge k)(x)\, dx = \int_A \big((f \wedge k)1_{B_k}\big)(x)\, dx \leqslant \int_A f(x)\, dx$, by the sandwich lemma we conclude that

$$\int_A f(x)\, dx = \lim_{\ell\to\infty} \int_{A\cap B(0,\ell)} (f \wedge \ell)(x)\, dx = \lim_{k\to\infty} \int_{A\cap B_k} (f \wedge k)(x)\, dx\,. \qquad \square$$

In other words, as long as $\{B_k\}_{k=1}^\infty$ "expands to the whole space", one can evaluate the improper integral using

$$\int_A f(x)\, dx = \lim_{k\to\infty} \int_{A\cap B_k} (f \wedge k)(x)\, dx\,.$$

One particular sequence of sets $\{B_k\}_{k=1}^\infty$ is given by $B_k = [-k,k] \times \cdots \times [-k,k]$.

Example 6.89. Consider the improper integral $\displaystyle\int_{-\infty}^{\infty} e^{-x^2}\, dx$. Instead of evaluating this improper integral directly, we consider the improper integral $\displaystyle\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\, d\mathbb{A}$. Note that Theorem 6.88 implies that

$$\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\, d\mathbb{A} = \lim_{k\to\infty} \int_{[-k,k]\times[-k,k]} e^{-(x^2+y^2)}\, d\mathbb{A} = \lim_{k\to\infty} \int_{B(0,k)} e^{-(x^2+y^2)}\, d\mathbb{A}\,.$$

By the Fubini theorem,

$$\lim_{k\to\infty} \int_{[-k,k]\times[-k,k]} e^{-(x^2+y^2)}\, d\mathbb{A} = \lim_{k\to\infty} \int_{-k}^{k} \Big(\int_{-k}^{k} e^{-(x^2+y^2)} dy \Big) dx = \Big(\int_{-\infty}^{\infty} e^{-x^2} dx \Big)^2,$$

while the change of variables formula (with $(x,y) = (r\cos\theta, r\sin\theta)$) implies that

$$\begin{aligned}\lim_{k\to\infty} \int_{B(0,k)} e^{-(x^2+y^2)}\, d\mathbb{A} &= \lim_{k\to\infty} \int_{[0,k]\times[0,2\pi]} e^{-r^2} r\, d(r,\theta) = \lim_{k\to\infty} \int_0^{2\pi} \Big(\int_0^k e^{-r^2} r\, dr \Big) d\theta \\ &= \lim_{k\to\infty} \int_0^{2\pi} \Big(\frac{e^{-r^2}}{-2} \Big|_{r=0}^{r=k} \Big) d\theta = \lim_{k\to\infty} \pi(1 - e^{-k^2}) = \pi\,.\end{aligned}$$

Since $\displaystyle\int_{-\infty}^{\infty} e^{-x^2}\, dx \geqslant 0$, we must have

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}\,.$$

Now we define the improper integrals for general functions. Let $\vee$ be an operation which outputs the maximum of values from both sides of $\vee$; that is,

$$(f \vee g)(x) = \max\big\{f(x), g(x)\big\}\,.$$

Define the positive and negative parts, denoted by f^+ and f^- respective, by $f^+ = f \vee 0$ and $f^- = (-f)\vee 0$. Since f^+, f^- are non-negative and $f = f^+ - f^-$, to defined the integral of f it is natural to consider the difference of the integrals $\int_A f^+(x)\,dx$ and $\int_A f^-(x)\,dx$. Note that if the collection of discontinuities of f has measure zero, the collections of discontinuities of both f^+ and f^- are sets of measure zero; thus the integrals $\int_A f^\pm(x)\,dx$ makes sense.

The discussion above motivates the following

Definition 6.90. Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f : A \to \mathbb{R}$ be a Riemann measurable function. f is said to be ***integrable*** on A if both integrals

$$\int_A f^+(x)\,dx \qquad \text{and} \qquad \int_A f^-(x)\,dx$$

are finite, where f^+ and f^- are the positive and negative parts of f defined by

$$f^+ = f \vee 0 \quad \text{and} \quad f^- = (-f) \vee 0\,.$$

If f is integrable on A, the integral of f on A, denoted by $\int_A f(x)\,dx$, is the number $\int_A f^+(x)\,dx - \int_A f^-(x)\,dx$.

Remark 6.91. 1. If f is integrable on A, then

$$\int_A \big|f(x)\big|\,dx = \int_A f^+(x)\,dx + \int_A f^-(x)\,dx < \infty;$$

thus the integrability of f on A sometimes is also called the absolute integrability of f on A or that the integral $\int_A f(x)\,dx$ is absolutely convergent.

2. For integrable function $f : A \to \mathbb{R}$, one can compute the integral of f on A by

$$\begin{aligned}\int_A f(x)\,dx &= \lim_{k\to\infty}\int_{A\cap B(0,k)} (f^+ \wedge k)(x)\,dx - \lim_{k\to\infty}\int_{A\cap B(0,k)} (f^- \wedge k)(x)\,dx \\ &= \lim_{k\to\infty}\int_{A\cap B(0,k)} f^+(x)\,dx - \lim_{k\to\infty}\int_{A\cap B(0,k)} f^-(x)\,dx \\ &= \lim_{k\to\infty}\int_{A\cap B(0,k)} \big(f^+(x) - f^-(x)\big)\,dx = \lim_{k\to\infty}\int_{A\cap B(0,k)} f(x)\,dx\,. \qquad (6.6.9)\end{aligned}$$

where (6.6.3) is used to conclude the third equality. In (6.6.9), the set $B(0,k)$ can also be replaced by increasing sequence of set $\{B_k\}_{k=1}^{\infty}$ as introduced in Theorem 6.88.

By Proposition 6.78, 6.81, 6.82 and Corollary 6.80, we can also establish the following

Theorem 6.92. *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f, g : A \to \mathbb{R}$ be integrable functions. If $f \leqslant g$, then*
$$\int_A f(x)\,dx \leqslant \int_A g(x)\,dx\,.$$

Theorem 6.93. *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f : A \to \mathbb{R}$ be an integrable function. Then for all $\alpha \in \mathbb{R}$,*
$$\int_A (\alpha f)(x)\,dx = \alpha \int_A f(x)\,dx\,.$$

Theorem 6.94. *Let $A, B \subseteq \mathbb{R}^n$ be Riemann measurable sets, and $f : A \cup B \to \mathbb{R}$ be an integrable function. If $A \cap B$ has measure zero, then*
$$\int_{A\cup B} f(x)\,dx = \int_A f(x)\,dx + \int_B f(x)\,dx\,.$$

Theorem 6.95. *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f, g : A \to \mathbb{R}$ be integrable functions. Then*
$$\int_A (f+g)(x)\,dx = \int_A f(x)\,dx + \int_A g(x)\,dx\,.$$

The proofs for the theorems above are left to the readers as exercises.

Remark 6.96. 1. If f is not integrable on A but one of the integrals $\int_A f^+(x)\,dx$ or $\int_A f^-(x)\,dx$ is finite, the number $\int_A f^+(x)\,dx - \int_A f^-(x)\,dx$ is still well-understood, and we still call this difference as the integral of f on A.

2. When at least one of the integrals $\int_A f^+(x)\,dx$ or $\int_A f^-(x)\,dx$ is finite,
$$\begin{aligned}
\int_A f(x)\,dx &= \lim_{k\to\infty} \int_{A\cap B(0,k)} (f^+ \wedge k)(x)\,dx - \lim_{k\to\infty} \int_{A\cap B(0,k)} (f^- \wedge k)(x)\,dx \\
&= \lim_{k\to\infty} \Big(\int_{A\cap B(0,k)} (f^+ \wedge k)(x)\,dx - \int_{A\cap B(0,k)} (f^- \wedge k)(x)\,dx \Big) \\
&= \lim_{k\to\infty} \int_{A\cap B(0,k)} (-k) \vee (f \wedge k)(x)\,dx\,.
\end{aligned}$$

Therefore, it is tempting to define the integrability of f on A by the existence of the limit

$$\lim_{k\to\infty}\int_{A\cap B(0,k)}\big[(-k)\vee(f\wedge k)\big](x)\,dx\,.$$

However, this cannot be the correct definition since if we adapt this definition, then the function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=\dfrac{1}{x}$ for $x\neq 0$ (and $f(0)$ is given arbitrarily) will be integrable on $\mathbb{R}$ (and the integral is 0 by symmetry), while Theorem 6.94, a should-have theorem for integrable functions, fails to hold for this particular function since

$$0=\int_{\mathbb{R}}f(x)\,dx\neq\int_{[0,\infty)}f(x)\,dx+\int_{(-\infty,0]}f(x)\,dx\,.$$

Theorem 6.97 (Comparison Test)**.** *Let $A\subseteq\mathbb{R}^n$ be a Riemann measurable set, $f,g:A\to\mathbb{R}$ be Riemann measurable functions. If $|f|\leqslant g$ on A and g is integrable on A, then f is integrable on A.*

Proof. Since $|f|=f^++f^-$, the condition that $|f|\leqslant g$ implies that $f^+\leqslant g$ and $f^-\leqslant g$; thus

$$\int_{A\cap B(0,k)}(f^\pm\wedge k)(x)\,dx\leqslant\int_{A\cap B(0,k)}(g\wedge k)(x)\,dx\leqslant\int_A g(x)\,dx<\infty\,.$$

Since $\displaystyle\int_{A\cap B(0,k)}(f^\pm\wedge k)(x)\,dx$ are increasing in k, both limits

$$\lim_{k\to\infty}\int_{A\cap B(0,k)}(f^\pm\wedge k)(x)\,dx$$

must exist (and are finite). Therefore, f is integrable on A. □

Example 6.98. Let $f:[0,\infty)\to\mathbb{R}$ be given by $f(x)=\dfrac{\sin x}{x^2+1}$. Then $\big|f(x)\big|\leqslant\dfrac{1}{x^2+1}$ and the function $y=\dfrac{1}{x^2+1}$ is integrable on $[0,\infty)$ since

$$\lim_{R\to\infty}\int_0^R\frac{1}{x^2+1}dx=\lim_{R\to\infty}\tan^{-1}x\Big|_{x=0}^{x=R}=\lim_{R\to\infty}\tan^{-1}R=\frac{\pi}{2}\,.$$

Let $\{a_k\}_{k=1}^\infty$ be a non-negative sequence. Then the series $\sum\limits_{k=1}^\infty a_k$ can be viewed as the integral of the piecewise constant function

$$f(x)=a_k\quad\text{if}\quad k-1\leqslant x<k \tag{6.6.10}$$

over the set $\mathbb{R}^+ \equiv \{x \geqslant 0\}$. Now suppose that $\{a_k\}_{k=1}^\infty$ is a general sequence in $\mathbb{R}$. Define $a_k^+ = \max\{a_k, 0\}$, $a_k^- = \max\{-a_k, 0\}$ for each $k \in \mathbb{N}$, and let $f : \mathbb{R}^+ \to \mathbb{R}$ be defined by (6.6.10). Recall that a series $\sum\limits_{k=1}^\infty a_k$ is absolutely convergent if $\sum\limits_{k=1}^\infty |a_k| < \infty$ and this is equivalent to that

$$\sum_{k=1}^\infty a_k^+ < \infty \quad \text{and} \quad \sum_{k=1}^\infty a_k^- < \infty .$$

Since $\displaystyle\int_{\mathbb{R}^+} f^+(x)\,dx = \sum_{k=1}^\infty a_k^+$ and $\displaystyle\int_{\mathbb{R}^+} f^-(x)\,dx = \sum_{k=1}^\infty a_k^-$, we find that the series $\sum\limits_{k=1}^\infty a_k$ is absolutely convergent if and only if f given by (6.6.10) is integrable on $\mathbb{R}^+$.

There is another concept of convergence of series, called the conditional convergence. Recall that a series $\sum\limits_{k=1}^\infty a_k$ is said to be conditionally convergent if the limit $\lim\limits_{\ell\to\infty} \sum\limits_{k=1}^{\ell} a_k$ exists but $\sum\limits_{k=1}^\infty |a_k| = \infty$. Let $\sum\limits_{k=1}^\infty a_k$ be a conditionally convergent series, and $f : \mathbb{R}^+ \to \mathbb{R}$ be given by (6.6.10). Then

$$\int_{\mathbb{R}^+} f^+(x)\,dx = \sum_{k=1}^\infty a_k^+ = \infty \quad \text{and} \quad \int_{\mathbb{R}^+} f^-(x)\,dx = \sum_{k=1}^\infty a_k^- = \infty ,$$

while the limit $\displaystyle\lim_{\ell\to\infty} \int_0^\ell f(x)\,dx = \lim_{\ell\to\infty} \sum_{k=1}^{\ell} a_k$ exists. The connection between the two kinds of convergence of series and integrals motivates the following

Definition 6.99. Let $A \subseteq \mathbb{R}$ be a Riemann measurable set, and $f : A \to \mathbb{R}$ be a Riemann measurable function. The improper integral $\displaystyle\int_A f(x)\,dx$ is said to be ***conditionally convergent*** if f is <u>not</u> integrable on A but the limit $\displaystyle\lim_{k,\ell\to\infty} \int_{A\cap(-\ell,k)} f(x)\,dx$ exists.

Remark 6.100. Suppose that the series $\sum\limits_{k=1}^\infty a_k$ is conditionally convergent. Then for each $r \in \mathbb{R}$, there exists a permutation $\pi : \mathbb{N} \to \mathbb{N}$ such that

$$r = \sum_{k=1}^\infty a_{\pi(k)} = a_{\pi(1)} + a_{\pi(2)} + a_{\pi(3)} + \cdots .$$

In other words, the order of the summation matters in a conditional convergent series; thus in general it will not be possible to talk about conditionally convergent integrals of functions on subsets of $\mathbb{R}^n$ if $n \geqslant 2$ since it usually requires the Fubini theorem to evaluate the multiple integrals while the Fubini theorem involves evaluating integrals in different orders.

Example 6.101. Let $f : [0,\infty) \to \mathbb{R}$ be given by $f(x) = \dfrac{\sin x}{x}$. Then for all $0 < r < R$, using the integration by parts formula we obtain that

$$\Big|\int_r^R \frac{\sin x}{x}dx\Big| = \Big|\frac{1-\cos x}{x}\Big|_{x=r}^{x=R} + \int_r^R \frac{1-\cos x}{x^2}dx\Big| \leqslant \frac{2}{R} + \frac{2}{r} + \int_r^R \frac{2}{x^2}dx = \frac{4}{r}.$$

Let $I_k = \displaystyle\int_0^k \frac{\sin x}{x}dx$. Then the inequality above implies that $\{I_k\}_{k=1}^\infty$ is Cauchy in $\mathbb{R}$, so the limit $\displaystyle\int_0^\infty \frac{\sin x}{x}dx = \lim_{k\to\infty} I_k$ exists. However,

$$\int_0^\infty f^+(x)\,dx = \sum_{k=1}^\infty \int_{(2k-2)\pi}^{(2k-1)\pi} \frac{\sin x}{x}dx \geqslant \sum_{k=1}^\infty \frac{1}{(2k-1)\pi}\int_{(2k-2)\pi}^{(2k-1)\pi} \sin x dx = \frac{2}{\pi}\sum_{k=1}^\infty \frac{1}{2k-1} = \infty$$

and

$$\int_0^\infty f^-(x)\,dx = \sum_{k=1}^\infty \int_{(2k-1)\pi}^{2k\pi} \frac{-\sin x}{x}dx \geqslant \sum_{k=1}^\infty \frac{1}{2k\pi}\int_{(2k-1)\pi}^{2k\pi} (-\sin x)dx = \frac{1}{\pi}\sum_{k=1}^\infty \frac{1}{k} = \infty\,.$$

Therefore, the improper integral $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx$ is conditionally convergent.

6.6.1 The monotone convergence theorem and the dominated convergence theorem

In the remaining part of this section, we present some important theorems introduced in Section 6.4 under the new settings of improper integrals.

Theorem 6.102 (Dominated Convergence Theorem)**.** *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f_k, f : A \to \mathbb{R}$ be Riemann measurable functions such that $\lim\limits_{k\to\infty} f_k(x) = f(x)$ for all $x \in A$. Suppose that there exists an integrable function g such that $|f_k| \leqslant g$ for all $k \in \mathbb{N}$. Then f is integrable on A, and*

$$\int_A f(x)\,dx = \lim_{k\to\infty}\int_A f_k(x)\,dx\,.$$

Proof. Since $\big|f_k(x)\big| \leqslant g(x)$ for all $x \in A$ and $k \in \mathbb{N}$, $\big|f(x)\big| \leqslant g(x)$ for all $x \in A$. By the integrability of g, the comparison test (Theorem 6.97) implies that f_k and f are also integrable on A.

Let $\varepsilon > 0$ be given. Since f, g are integrable on A, there exists $L > 0$ such that

$$0 \leqslant \int_A g(x)\,dx - \int_{A\cap B(0,\ell)} (g\wedge \ell)(x)\,dx < \frac{\varepsilon}{3} \qquad \forall\, \ell \geqslant L \tag{6.6.11}$$

and

$$\Big|\int_A f(x)\,dx - \int_{A\cap B(0,\ell)} \big((-\ell)\vee(f\wedge \ell)\big)(x)\,dx\Big| < \frac{\varepsilon}{3} \qquad \forall\,\ell \geqslant L\,.$$

Moreover, since $(-L)\vee(f_k\wedge L) \to (-L)\vee(f\wedge L)$ p.w. as $k\to\infty$ (due to the pointwise convergence of $\{f_k\}_{n=1}^{\infty}$ to f), and $\big|(-L)\vee(f_k\wedge L)\big| \leqslant L$ on $A\cap B(0,L)$, the Bounded Convergence Theorem (Theorem 6.70) implies that there exists $K>0$ such that

$$\Big|\int_{A\cap B(0,L)} \big((-L)\vee(f_k\wedge L)\big)(x)\,dx - \int_{A\cap B(0,L)} \big((-L)\vee(f\wedge L)\big)(x)\,dx\Big| < \frac{\varepsilon}{3} \quad \forall\,k\geqslant K\,.$$

Note that Theorem 6.81 implies that

$$\begin{aligned}\Big|\int_A f(x)\,dx - \int_A f_k(x)\,dx\Big| &\leqslant \Big|\int_A f(x)\,dx - \int_{A\cap B(0,L)} \big((-L)\vee(f\wedge L)\big)(x)\,dx\Big| \\ &\quad + \Big|\int_{A\cap B(0,L)} \big((-L)\vee(f\wedge L)\big)(x)\,dx - \int_{A\cap B(0,L)} \big((-L)\vee(f_k\wedge L)\big)(x)\,dx\Big| \\ &\quad + \Big|\int_{A\cap B(0,L)} \big((-L)\vee(f_k\wedge L)\big)(x)\,dx - \int_{A\cap B(0,L)} f_k(x)\,dx\Big| + \Big|\int_{A\cap B(0,L)^{\complement}} f_k(x)\,dx\Big|,\end{aligned}$$

and the fact that $|f_k|\leqslant g$ implies that

$$\big|\big((-L)\vee(f_k\wedge L)\big)(x) - f_k(x)\big| \leqslant g(x) - (g\wedge L)(x)\,.$$

Therefore, for $k\geqslant K$,

$$\begin{aligned}&\Big|\int_A f(x)\,dx - \int_A f_k(x)\,dx\Big| \\ &\qquad < \frac{2\varepsilon}{3} + \int_{A\cap B(0,L)} \big(g(x) - (g\wedge L)(x)\big)\,dx + \int_{A\cap B(0,L)^{\complement}} g(x)\,dx \\ &\qquad \leqslant \frac{2\varepsilon}{3} + \int_A g(x)\,dx - \int_{A\cap B(0,L)} (g\wedge L)(x)\,dx < \varepsilon\,.\end{aligned}$$

□

The Monotone Convergence Theorem for improper integrals, unlike the case in the Riemann integrals, is no longer an immediate consequence of the Dominated Convergence Theorem since the "integral" of the limit function might be infinite. It requires a little bit more attention to get proved.

Theorem 6.103 (Monotone Convergence Theorem for Improper Integrals)**.** *Let $A\subseteq\mathbb{R}^n$ be a Riemann measurable set, and $f_k, f : A\to\mathbb{R}$ be non-negative, Riemann measurable* *functions such that* $\lim\limits_{k\to\infty} f_k(x) = f(x)$ *for all* $x \in A$. *Suppose that* $\{f_k\}_{k=1}^\infty$ *is a monotone increasing sequence of functions; that is,* $f_k \leqslant f_{k+1}$ *for all* $k \in \mathbb{N}$. *Then*

$$\int_A f(x)\,dx = \lim_{k\to\infty} \int_A f_k(x)\,dx\,. \tag{6.6.12}$$

Proof. By the Dominated Convergence Theorem (Theorem 6.102), we only need to consider the case that $\displaystyle\int_A f(x)\,dx = \infty$ and show that $\displaystyle\lim_{k\to\infty}\int_A f_k(x)\,dx = \infty$. We also assume the non-trivial case that $\displaystyle\int_A f_k(x)\,dx < \infty$ for all $k \in \mathbb{N}$.

Let $M > 0$ be given. Since $\displaystyle\int_A f(x)\,dx = \infty$, there exists $L > 0$ such that

$$\int_{A\cap B(0,\ell)} (f \wedge \ell)(x)\,dx \geqslant 2M \qquad \forall\, \ell \geqslant L\,.$$

By the Monotone Convergence Theorem for Riemann integrals (Theorem 6.69), there exists $K > 0$ such that

$$-M \leqslant \int_{A\cap B(0,L)} (f_k \wedge L)(x)\,dx - \int_{A\cap B(0,L)} (f \wedge L)(x)\,dx \leqslant 0 \qquad \forall\, k \geqslant K\,.$$

Therefore, for all $k \geqslant K$,

$$\begin{aligned}
\int_A f_k(x)\,dx &= \int_A f_k(x)\,dx - \int_{A\cap B(0,L)} (f_k \wedge L)(x)\,dx + \int_{A\cap B(0,L)} (f_k \wedge L)(x)\,dx \\
&\qquad - \int_{A\cap B(0,L)} (f \wedge L)(x)\,dx + \int_{A\cap B(0,L)} (f \wedge L)(x)\,dx \\
&\geqslant \int_A f_k(x)\,dx - \int_{A\cap B(0,L)} (f_k \wedge L)(x)\,dx + M \geqslant M\,.
\end{aligned}$$

□

When non-negativity of functions is removed from the condition, for (6.6.12) to hold it is required that the sequence of functions has an integrable lower bound. To be more precise, we have the following

Corollary 6.104. *Let* $A \subseteq \mathbb{R}^n$ *be a Riemann measurable set, and* $f_k, f : A \to \mathbb{R}$ *be Riemann measurable functions such that* $\lim\limits_{k\to\infty} f_k(x) = f(x)$ *for all* $x \in A$. *Suppose that there exists an integrable function* $g : A \to \mathbb{R}$ *such that* $f_k(x) \geqslant g(x)$ *for all* $x \in A$ *and* $k \in \mathbb{N}$, *and* $\{f_k\}_{k=1}^\infty$ *is a monotone increasing sequence of functions; that is,* $f_k \leqslant f_{k+1}$ *for all* $k \in \mathbb{N}$. *Then*

$$\int_A f(x)\,dx = \lim_{k\to\infty} \int_A f_k(x)\,dx\,.$$

Proof. Consider the new sequence of functions $\{h_k\}_{k=1}^{\infty}$ defined by $h_k = f_k - g$ and apply the Monotone Convergence Theorem (Theorem 6.103). $\square$

For monotone decreasing sequences of functions, we have the following

Corollary 6.105. *Let $A \subseteq \mathbb{R}^n$ be a Riemann measurable set, and $f_k, f : A \to \mathbb{R}$ be Riemann measurable functions such that $\lim\limits_{k\to\infty} f_k(x) = f(x)$ for all $x \in A$. Suppose that f_1 is integrable on A and $\{f_k\}_{k=1}^{\infty}$ is a monotone decreasing sequence of functions; that is, $f_k \geqslant f_{k+1}$ for all $k \in \mathbb{N}$. Then*

$$\int_A f(x)\,dx = \lim_{k\to\infty} \int_A f_k(x)\,dx\,.$$

Proof. Consider the new sequence of functions $\{h_k\}_{k=1}^{\infty}$ defined by $h_k = f_1 - f_k$ and apply the Monotone Convergence Theorem (Theorem 6.103). $\square$

6.6.2 The Fubini theorem and the Tonelli theorem

In this section we present the Fubini theorem for improper integrals. The Fubini theorem for improper integrals takes the form

$$\int_{A\times B} f(x,y)d(x,y) = \int_A \Big(\int_B f(x,y)dy\Big)dx$$

or

$$\int_{A\times B} f(x,y)d(x,y) = \int_B \Big(\int_A f(x,y)dx\Big)dy\,.$$

as long as f satisfies certain conditions. However, the iterated integrals on the right-hand side will be meaningless if the functions $F(x) \equiv \displaystyle\int_B f(\cdot,y)dy$ and $G(y) \equiv \displaystyle\int_A f(x,\cdot)\,dx$ are not Riemann measurable; thus in general we need to impose the condition that F and G are Riemann measurable. In fact, even if $f : A\times B \to \mathbb{R}$ is continuous, F might still be discontinuous at some points. For example, let $A = [-1,1]$, $B = [1,\infty)$ and $f(x,y) = |x|y^{-1-|x|}$. Then

$$F(x) = \begin{cases} 0 & \text{if } x = 0\,,\\ 1 & \text{otherwise}\,,\end{cases}$$

which is discontinuous at $x = 0$. In other words, the collection of discontinuities of F might not be empty even if f is continuous on $A \times B$ since "partial integration" might produce extra discontinuities; thus in general we do not know if F is Riemann measurable even if f is continuous.

Before proceeding to the Fubini theorem, we first establish the following

Theorem 6.106 (Tonelli). *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be Riemann measurable sets such that $A \times B$ is Riemann measurable, and $f : A \times B \to \mathbb{R}$ be a non-negative Riemann measurable function.*

1. *If for all $x \in A$, $f(x, \cdot)$ is integrable on B and the function $\displaystyle\int_B f(\cdot, y)\, dy : A \to \mathbb{R}$ is Riemann measurable, then*
$$\int_{A\times B} f(x,y)d(x,y) = \int_A \Big(\int_B f(x,y)dy\Big)dx\,.$$

2. *If for all $y \in B$, $f(\cdot, y)$ is integrable on A and the function $\displaystyle\int_A f(x, \cdot)\, dx : B \to \mathbb{R}$ is Riemann measurable, then*
$$\int_{A\times B} f(x,y)d(x,y) = \int_B \Big(\int_A f(x,y)dx\Big)dy\,.$$

Proof. It suffices to show the first case since the proof of the other case is similar.

We show that for each $k \in \mathbb{N}$, the function $g_k : A \to \mathbb{R}$ defined by
$$g_k(x) = \int_{B\cap[-k,k]^m} (f \wedge k)(x,y)dy$$
is Riemann measurable. We note that the fact that $f(x, \cdot)$ is integrable on B for each $x \in A$ implies that g_k is well-defined for all $x \in A$.

For each $\ell \in \mathbb{N}$, the Fubini theorem for Riemann integrals provides that
$$\int_{(A\cap[-\ell,\ell]^n)\times(B\cap[-k,k]^m)} (f\wedge k)(x,y)d(x,y) = \underline{\int}_{A\cap[-\ell,\ell]^n} g_k(x)\, dx = \overline{\int}_{A\cap[-\ell,\ell]^n} g_k(x)\, dx\,.$$
Therefore, g_k is Riemann integrable on $A \cap [-\ell, \ell]^n$ for all $\ell \in \mathbb{N}$, and the Lebesgue theorem (Theorem 6.32) implies that the collection of discontinuities of g_k in $A \cap [-\ell, \ell]^n$ has measure zero. By Theorem 6.26 and the fact that $A = \bigcup\limits_{\ell=1}^{\infty} (A \cap [-\ell, \ell]^n)$, the collection of discontinuities of g_k in A has measure zero; thus g_k is Riemann measurable.

For each $k \in \mathbb{N}$, define $f_k(x, y) = \mathbf{1}_{[-k,k]^{n+m}}(x, y)(f \wedge k)(x, y)$ and $h_k(x) = \mathbf{1}_{[-k,k]^n}(x)g_k(x)$. Then $\{f_k\}_{k=1}^{\infty}$ and $\{h_k\}_{k=1}^{\infty}$ are non-negative monotone increasing sequences. Moreover, it is clear that $\{f_k\}_{k=1}^{\infty}$ converges pointwise to f (on A) and $\{h_k\}_{k=1}^{\infty}$ converges pointwise to the function $\displaystyle\int_B f(\cdot,y)\,dy$. By the Fubini theorem for Riemann integrals again,

$$\begin{aligned}\int_{A\times B} f_k(x,y)d(x,y) &= \int_{(A\cap[-k,k]^n)\times(B\cap[-k,k]^m)} (f\wedge k)(x,y)d(x,y)\\ &= \int_{A\cap[-k,k]^n} g_k(x)\,dx = \int_A h_k(x)\,dx\,;\end{aligned}$$

thus the Monotone Convergence Theorem (Theorem 6.103) implies that

$$\begin{aligned}\int_{A\times B} f(x,y)d(x,y) &= \lim_{k\to\infty}\int_{A\times B} f_k(x,y)d(x,y) = \lim_{k\to\infty}\int_A h_k(x)\,dx\\ &= \int_A\Big(\int_B f(x,y)dy\Big)dx\,.\end{aligned}$$

□

Now we can present the Fubini theorem.

Theorem 6.107 (Fubini). *Let $A\subseteq\mathbb{R}^n$ and $B\subseteq\mathbb{R}^m$ be Riemann measurable sets such that $A\times B$ is Riemann measurable, and $f:A\times B\to\mathbb{R}$ be an integrable function.*

1. *If for all $x\in A$, $f(x,\cdot)$ is integrable on B, and the functions $\displaystyle\int_B f(\cdot,y)dy: A\to\mathbb{R}$ and $\displaystyle\int_B \big|f(\cdot,y)\big|dy: A\to\mathbb{R}$ are Riemann measurable, then*
$$\int_{A\times B} f(x,y)d(x,y) = \int_A\Big(\int_B f(x,y)dy\Big)dx\,.$$

2. *If for all $y\in B$, $f(\cdot,y)$ is integrable on A, and the functions $\displaystyle\int_A f(x,\cdot)dx: B\to\mathbb{R}$ and $\displaystyle\int_A \big|f(x,\cdot)\big|dx: B\to\mathbb{R}$ are Riemann measurable, then*
$$\int_{A\times B} f(x,y)d(x,y) = \int_B\Big(\int_A f(x,y)dx\Big)dy\,.$$

Proof. It suffices to prove the first case.

Since $f^\pm = \dfrac{1}{2}(|f|+f)$, by assumption the functions $\displaystyle\int_B f^+(\cdot,y)dy : A\to\mathbb{R}$ and $\displaystyle\int_B f^-(\cdot,y)dy : A\to\mathbb{R}$ are Riemann measurable functions. Therefore, the Tonelli theorem implies that

$$\int_{A\times B} f^\pm(x,y)d(x,y) = \int_A\Big(\int_B f^\pm(x,y)\,dy\Big)dx\,;$$

thus

$$\begin{aligned}\int_{A\times B} f(x,y)d(x,y) &= \int_{A\times B} f^+(x,y)d(x,y) - \int_{A\times B} f^-(x,y)d(x,y)\\ &= \int_A \Big(\int_B f^+(x,y)\,dy\Big)dx - \int_A \Big(\int_B f^-(x,y)\,dy\Big)dx\\ &= \int_A \Big(\int_B f^+(x,y)\,dy - \int_A\int_B f^-(x,y)\,dy\Big)dx = \int_A \Big(\int_B f(x,y)\,dy\Big)dx\,. \end{aligned}$$ ▫

Corollary 6.108. *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be Riemann measurable sets such that $A \times B$ is Riemann measurable, and $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ be integrable functions. Then the function $h : A \times B \to \mathbb{R}$ given by $h(x, y) = f(x)g(y)$ is integrable, and*

$$\int_{A\times B} h(x,y)d(x,y) = \Big(\int_A f(x)\,dx\Big)\Big(\int_B g(y)\,dy\Big).$$

Proof. By Theorem 6.28, $|h|$ is Riemann measurable. Moreover, by the integrability of g we find that for each $x \in A$ the function $\big|h(x,\cdot)\big| : B \to \mathbb{R}$ is Riemann measurable. Since $|f|$ is integrable on A and

$$\int_B \big|h(x,y)\big|\,dy = \big|f(x)\big| \int_B \big|g(y)\big|\,dy\,,$$

the function $\displaystyle\int_B |h(\cdot,y)|dy : A \to \mathbb{R}$ is Riemann measurable. In other words, h satisfies conditions in the Tonelli theorem (Theorem 6.106); thus we have

$$\int_{A\times B} \big|h(x,y)\big|d(x,y) = \int_A \Big(\int_B \big|h(x,y)\big|dy\Big)dx = \Big(\int_A \big|f(x)\big|\,dx\Big)\Big(\int_B \big|g(y)\big|\,dy\Big) < \infty\,.$$

Therefore, h is integrable on $A \times B$. Since $h(x,\cdot)$ is integrable on B for all $x \in A$ and $h(\cdot,y)$ is integrable on A for all $y \in B$, the Fubini theorem (Theorem 6.107) further implies that

$$\int_{A\times B} h(x,y)d(x,y) = \int_A \Big(\int_B h(x,y)dy\Big)dx = \Big(\int_A f(x)\,dx\Big)\Big(\int_B g(y)\,dy\Big).$$ ▫

6.7 Exercises

§ 6.1 Integrable Functions

Problem 6.1. Let $f : [a,b] \to \mathbb{R}$ be a bounded function, and $\mathcal{P}_n$ denote the division of $[a,b]$ into 2^n equal sub-intervals. Show that f is Riemann integrable on $[a,b]$ if and only if

$$\lim_{n\to\infty} U(f,\mathcal{P}_n) = \lim_{n\to\infty} L(f,\mathcal{P}_n)\,.$$

Problem 6.2. Let $f, g : [a, b] \to \mathbb{R}$ be functions, where g is continuous, and f be non-negative, bounded, Riemann integrable on $[a, b]$. Show that fg is Riemann integrable.

Problem 6.3. Let $f : [a, b] \to \mathbb{R}$ be differentiable and assume that f' is Riemann integrable. Prove that $\displaystyle\int_a^b f'(x)\, dx = f(b) - f(a)$.

Hint: Use the Mean Value Theorem.

Problem 6.4. Suppose that $f : [a, b] \to \mathbb{R}$ is Riemann integrable, $m \leqslant f(x) \leqslant M$ for all $x \in [a, b]$, and $\varphi : [m, M] \to \mathbb{R}$ is continuous. Show that $\varphi \circ f$ is Riemann integrable on $[a, b]$.

Problem 6.5. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f, g : A \to \mathbb{R}$ be functions. Show that

$$\underline{\int_A} f(x)\, dx \leqslant \underline{\int_A} g(x)\, dx \qquad \text{and} \qquad \overline{\int_A} f(x)\, dx \leqslant \overline{\int_A} g(x)\, dx\,.$$

Problem 6.6. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : [a, b] \to \mathbb{R}$ be a function.

1. Show that if f is Riemann integrable on A, then f is bounded.

2. Show that if f is Darboux integrable on A, then f is bounded.

Note that we in some sense use these properties in the proof of the equivalence between the Riemann integrability and the Darboux integrability, so you'd better no use this equivalence in the proof.

Hint:

1. Let $\mathrm{I} = (R)\displaystyle\int_A f(x)\, dx$. There exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of A, then any Riemann sum of f for $\mathcal{P}$ locates in $(\mathrm{I}-1, \mathrm{I}+1)$. Let $\mathcal{P} = \{\Delta_1, \Delta_2, \cdots, \Delta_N\}$ be a partition of A satisfying $\|\mathcal{P}\| < \delta$, and for each $1 \leqslant k \leqslant N$, let c_k be the center of Δ_k. Show that for each $1 \leqslant \ell \leqslant N$,
$$\begin{aligned}\frac{1}{\nu(\Delta_\ell)}\Big[\mathrm{I} - 1 - \sum_{1 \leqslant \ell \leqslant N, k \neq \ell} \overline{f}^A(c_k)\nu(\Delta_k)\Big] &< \overline{f}^A(x) \\ &< \frac{1}{\nu(\Delta_k)}\Big[\mathrm{I} + 1 - \sum_{1 \leqslant \ell \leqslant N, k \neq \ell} \overline{f}^A(c_k)\nu(\Delta_k)\Big] \qquad \forall\, x \in \Delta_\ell\,.\end{aligned}$$

2. Note that if $\mathcal{P}$ be a partition of A and $\Delta \in \mathcal{P}$, then
$$-\infty < \sup_{x\in\Delta} \overline{f}^A(x) \leqslant \infty \qquad \text{and} \qquad -\infty \leqslant \inf_{x\in\Delta} \overline{f}^A(x) < \infty\,.$$
What happened to $U(f, \mathcal{P})$ and $L(f, \mathcal{P})$ if f is not bounded?

Problem 6.7. Prove that $\lim_{n\to\infty} \frac{(n!)^{\frac{1}{n}}}{n} = e^{-1}$ by considering the upper and lower sums of $\int_{\frac{1}{n}}^{1} \log x \, dx$ for partition $\left\{\frac{1}{n}, \frac{2}{n}, \cdots, 1\right\}$.

Problem 6.8. Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be a bounded function such that $f(x,y) \leqslant f(x,z)$ if $y < z$ and $f(x,y) \leqslant f(t,z)$ if $x < t$. In other words, $f(x,\cdot)$ and $f(\cdot,y)$ are both non-decreasing functions for fixed $x, y \in [0,1]$. Show that f is Riemann integrable on $[0,1] \times [0,1]$.

Problem 6.9. For a function $f : [a,b] \to \mathbb{R}$, define the ***total variation*** of f on $[a,b]$ by

$$V_a^b(f) = \sup\left\{ \sum_{k=1}^{n} \left|f(x_k) - f(x_{k-1})\right| \,\middle|\, \{a = x_0 < \cdots < x_n = b\} \text{ is a partition of } [a,b] \right\}.$$

Sometimes $V_a^b(f)$ is written as $\|f\|_{\mathrm{TV}([a,b])}$.

A function $f : [a,b]$ is said to have ***bounded variation*** on $[c,d]$ or be ***of bounded variation*** on $[c,d]$, where $[c,d] \subseteq [a,b]$, if $V_c^d(f) < \infty$. Complete the following.

1. Let $\mathrm{BV}([a,b]) = \left\{f : [a,b] \to \mathbb{R} \,\middle|\, V_a^b(f) < \infty\right\}$, called the space of functions of bounded variation (on $[a,b]$). Show that $\mathrm{BV}([a,b])$ is a vector space.

2. Is V_a^b a norm on $\mathrm{BV}([a,b])$; that is, does $\|\cdot\|_{\mathrm{TV}([a,b])} : \mathrm{BV}([a,b]) \to \mathbb{R}$ defined by $\|f\| \equiv V_a^b(f)$ satisfy Definition 2.15?

3. Recall that $\mathscr{C}^1([a,b];\mathbb{R}) \equiv \left\{f : [a,b] \to \mathbb{R} \,\middle|\, f' \text{ is continuous on } [a,b]\right\}$. Show that if $f \in \mathscr{C}^1([a,b];\mathbb{R})$, then f is of bounded variation.

4. Show that if $f \in \mathscr{C}^1([a,b];\mathbb{R})$, then $V_a^b(f) = \int_a^b \left|f'(x)\right| dx$.

5. Show that if $V_a^b(f) < \infty$ (f is not necessarily differentiable everywhere), then

$$V_a^b(f) = \sup\left\{ \int_a^b f(x)\phi'(x)\,dx \,\middle|\, \phi \in \mathscr{C}^1([a,b];\mathbb{R}), |\phi(x)| \leqslant 1 \text{ for all } x \in [a,b], \phi(a) = \phi(b) = 0 \right\}.$$

§ 6.2 Lebesgue's Theorem

Problem 6.10. Complete the following.

1. Show that if A is a set of volume zero, then A has measure zero. Is it true that if A has measure zero, then A also has volume zero?

2. Let $a, b \in \mathbb{R}$ and $a < b$. Show that the interval $[a, b]$ does not have measure zero (in $\mathbb{R}$).

3. Let $A \subseteq [a, b]$ be a set of measure zero (in $\mathbb{R}$). Show that $[a, b] \backslash A$ does not have measure zero (in $\mathbb{R}$).

4. Show that the Cantor set (defined in Exercise Problem 3.19) has volume zero.

Problem 6.11. Let $A = \bigcup\limits_{k=1}^{\infty} B\big(\frac{1}{k}, \frac{1}{2^k}\big) = \bigcup\limits_{k=1}^{\infty} \big(\frac{1}{k} - \frac{1}{2^k}, \frac{1}{k} + \frac{1}{2^k}\big)$ be a subset of $\mathbb{R}$. Does A have volume?

Problem 6.12. Let $f : [a, b] \to \mathbb{R}$ be bounded and Riemann integrable. Show that the graph of f has volume zero by considering the difference of the upper and lower sums of f.

Problem 6.13. Let $A \subseteq \mathbb{R}^n$ be an open bounded set with volume, and $f : A \to \mathbb{R}$ be continuous. Show that if $\displaystyle\int_B f(x)\, dx = 0$ for all subsets $B \subseteq A$ with volume, then $f = 0$.

Problem 6.14. Prove the following statements.

1. The function $f(x) = \sin \dfrac{1}{x}$ is Riemann integrable on $(0, 1)$.

2. Let $f : (0, 1] \to \mathbb{R}$ be given by
$$f(x) = \begin{cases} \dfrac{1}{p} & \text{if } x = \dfrac{q}{p} \in \mathbb{Q},\ (p, q) = 1\,, \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$
Then f is Riemann integrable on $(0, 1]$. Find $\displaystyle\int_{(0,1]} f(x) dx$ as well.

3. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ is Riemann integrable. Then f^k（f 的 k 次方）is integrable for all $k \in \mathbb{N}$.

Problem 6.15. Suppose that $f : [a, b] \to \mathbb{R}$ is Riemann integrable, and the set $\big\{x \in [a, b] \,\big|\, f(x) \neq 0\big\}$ has measure zero. Show that $\displaystyle\int_a^b f(x)\, dx = 0$.

Problem 6.16. Let $f : [0, 1] \to \mathbb{R}$ be a function satisfying the property that for all $a, b \in (0, 1)$ with $a < b$, there exists $c \in (a, b)$ such that $f(c) = 1$ or $f(c) = 0$. Show that if f is Riemann integrable on $[0, 1]$ and $\displaystyle\int_0^1 f(x) dx \neq 1$, then the set $Z = \big\{x \in [0, 1] \,\big|\, f(x) = 0\big\}$ cannot have measure zero.

§ 6.3 Properties of the Integrals

Problem 6.17. Find an example that

$$\underline{\int_A} f(x)\,dx + \underline{\int_A} g(x)\,dx < \underline{\int_A} (f+g)(x)\,dx < \overline{\int_A} (f+g)(x)\,dx < \overline{\int_A} f(x)\,dx + \overline{\int_A} g(x)\,dx\,.$$

Problem 6.18. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. Show that if f is Riemann integrable on A, then $|f|$ is also Riemann integrable on A.

§ 6.4 Fubini's Theorem

Problem 6.19. 1. Evaluate the iterated integral $\displaystyle\int_0^1 \Big(\int_0^x (2y - y^2)^{\frac{2}{3}} dy\Big) dx$.

2. Draw the region corresponding to the integral $\displaystyle\int_0^1 \Big(\int_1^{e^x} (x+y)\,dy\Big) dx$ and evaluate.

3. Change the order of integration of the integral in 2 and check if the answer is unaltered.

Problem 6.20. Let $A = [a,b] \times [c,d]$ be a rectangle in $\mathbb{R}^2$, and $f : A \to \mathbb{R}$ be Riemann integrable. Show that the sets

$$\Big\{x \in [a,b] \,\Big|\, \underline{\int_c^d} f(x,y)dy \neq \overline{\int_c^d} f(x,y)dy\Big\} \quad \text{and} \quad \Big\{y \in [c,d] \,\Big|\, \underline{\int_a^b} f(x,y)dx \neq \overline{\int_a^b} f(x,y)dx\Big\}$$

have measure zero (in $\mathbb{R}^1$).

Problem 6.21. Define a set $S \subseteq [0,1] \times [0,1]$ by

$$S = \Big\{ \big(\frac{p}{m}, \frac{k}{m}\big) \in [0,1] \times [0,1] \,\Big|\, m, p, k \in \mathbb{N}\,, \gcd(m,p) = 1 \text{ and } 1 \leqslant k \leqslant m-1 \Big\}\,.$$

Show that

$$\int_0^1 \Big(\int_0^1 \mathbf{1}_S(x,y)\,dy\Big) dx = \int_0^1 \Big(\int_0^1 \mathbf{1}_S(x,y)\,dx\Big) dy = 0$$

but $\mathbf{1}_S$ is not Riemann integrable on $[0,1] \times [0,1]$.

Problem 6.22. Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be given by

$$f(x,y) = \begin{cases} 2^{2n} & \text{if } (x,y) \in [2^{-n}, 2^{-n+1}) \times [2^{-n}, 2^{-n+1}),\ n \in \mathbb{N}\,, \\ -2^{2n+1} & \text{if } (x,y) \in [2^{-n}, 2^{-n+1}) \times [2^{-n-1}, 2^{-n}),\ n \in \mathbb{N}\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

1. Show that $\displaystyle\int_0^1 f(x,y)\,dx = 0$ for all $y \in \big[0, \frac{1}{2}\big)$.
2. Show that $\displaystyle\int_0^1 f(x,y)\,dy = 0$ for all $x \in [0,1)$.
3. Justify if the iterated (improper) integrals $\displaystyle\int_0^1\int_0^1 f(x,y)\,dxdy$ and $\displaystyle\int_0^1\int_0^1 f(x,y)\,dydx$ are identical.

§ 6.5 The Monotone and Bounded Convergence Theorems

Problem 6.23. Let $A = [a,b]$ be a closed interval in $\mathbb{R}$, and $f_k : A \to \mathbb{R}$ be a non-decreasing sequence (that is, $f_k \leqslant f_{k+1}$ for all $k \in \mathbb{N}$) of continuous functions such that

$$\lim_{k\to\infty} f_k(x) = f(x) \qquad \forall\, x \in A$$

for some continuous function $f : A \to \mathbb{R}$. Show that

$$\lim_{k\to\infty} \sup_{x\in A} \big|f_k(x) - f(x)\big| = 0\,.$$

If we do not assume that f is continuous or if A is replaced by other kind of intervals, is the conclusion still true?
Hint: Mimic the proof of Lemma 6.64 in the lecture note.

Problem 6.24. Let $f : [a,b]\times[c,d] \to \mathbb{R}$ be a continuous function, and $F(x) = \displaystyle\int_c^d f(x,y)\,dy$. Use the bounded convergence theorem to show that F is continuous on $[a,b]$.

Problem 6.25 (The multiple integral version of Theorem 6.65)**.** Let A be a closed rectangle in $\mathbb{R}^n$, and $f_k : A \to \mathbb{R}$ be a decreasing sequence of bounded functions. Show (without applying Theorem 6.69 and 6.70) that if $\lim\limits_{k\to\infty} f_k(x) = 0$ for all $x \in A$, then

$$\lim_{k\to\infty} \underline{\int}_A f_k(x)\,dx = 0\,.$$

Conclude the Monotone Convergence Theorem (Theorem 6.69) and the Bounded Convergence Theorem (Theorem 6.70) using the this conclusion of convergence.

§ 6.6 Improper Integrals

Problem 6.26. Let $f : [2,\infty) \to \mathbb{R}$ be given by $f(x) = x^{-1}(\log x)^{-p}$. Show that f is integrable on $[2,\infty)$ if and only if $p > 1$.

Problem 6.27. Define the Beta function $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt$ whenever the integral makes sense.

1. Show that B is well-defined for $x, y > 0$.

2. Show that $B(x,y) = B(y,x)$.

3. Find $B(2,2)$ and $B(4,3)$.

Problem 6.28. Suppose that $f : (0,b] \to \mathbb{R}$ is continuous, positive, integrable on $(0,b]$, and that $f(x)$ increases monotonically to ∞ as x approaches 0 from the right. Show that $\lim\limits_{x\to 0^+} xf(x) = 0$.

Problem 6.29. Let $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$ be Riemann measurable sets, and $f : A \times B \to \mathbb{R}$ be non-negative, uniformly continuous and integrable on $A \times B$. Define $F(x) = \int_B f(x,y)\,dy$.

1. Show that if B is bounded, then $F : A \to \mathbb{R}$ is continuous. How about if B is not bounded?

2. Let f have the additional property that for each $\varepsilon > 0$, there exists $N > 0$ such that
$$\Big|\int_{B\cap B(0,k)} (f \wedge k)(x,y)\,dy - \int_B f(x,y)\,dy\Big| < \varepsilon \qquad \forall\, k \geqslant N \text{ and } x \in A\,.$$
Show that F is continuous on A. In particular, show that if $f(x,y) \leqslant g(y)$ for all $(x,y) \in A \times B$, and g is integrable on B, then F is continuous.

Problem 6.30. Let $f : \mathbb{R} \to \mathbb{R}$ be a Riemann measurable function, and $F : \mathbb{R} \to \mathbb{R}$ be defined by
$$F(x) = \int_{\mathbb{R}} f(y)\cos(x-y)\,dy$$
whenever the integral exists. Show that if the function f is integrable, then F is defined on $\mathbb{R}$ and is differentiable on $\mathbb{R}$ with derivative
$$F'(x) = \int_{\mathbb{R}} f(y)\frac{\partial}{\partial x}\cos(x-y)\,dy = -\int_{\mathbb{R}} f(y)\sin(x-y)\,dy\,.$$

Problem 6.31. Let $f : \mathbb{R} \to \mathbb{R}$ be an integrable Riemann measurable function, and $F : \mathbb{R} \to \mathbb{R}$ be defined by
$$F(x) = \int_{\mathbb{R}} f(y)\cos(xy)\,dy$$

(which exists for all $x \in \mathbb{R}$ since f is integrable). Show that if the function $g(x) = xf(x)$ is integrable, then F is differentiable on $\mathbb{R}$ and

$$F'(y) = \int_{\mathbb{R}} f(x) \frac{\partial}{\partial y} \cos(xy)\, dx = -\int_{\mathbb{R}} x f(x) \sin(xy)\, dx\,.$$

Problem 6.32. In each of the following problems, compute $F'(x)$.

1. $F(x) = \displaystyle\int_0^1 \frac{\sin(xy)}{1+y} dy$.
2. $F(x) = \displaystyle\int_1^2 \frac{e^{-y}}{1+xy} dy$.
3. $F(x) = \displaystyle\int_0^1 \frac{xy}{\sqrt{1-x^2y^2}} dy$, where $|x| < 1$.
4. $F(x) = \displaystyle\int_0^1 f(x,y)\, dy$, where $f(x,y) = \begin{cases} \dfrac{y^x - 1}{\log y} & \text{if } y \neq 0, 1\,, \\ 0 & \text{if } y = 0\,, \\ x & \text{if } y = 1\,. \end{cases}$
5. $F(x) = \displaystyle\int_0^\pi \log(1 - 2x\cos y + x^2) dy$.
6. $F(x) = \displaystyle\int_{\cos x}^{1+x^2} \frac{e^{-y}}{1+xy} dy$.
7. $F(x) = \displaystyle\int_0^{x^2} \tan^{-1}\frac{y}{x^2} dy$.

Problem 6.33. Show that if $m, n \in \mathbb{N}$, then

$$\int_0^1 t^n (\log t)^m dt = (-1)^m \frac{m!}{(n+1)^{m+1}}\,.$$

Hint: Differentiate $\displaystyle\int_0^1 x^n\, dx$ with respect to n and use induction.

Problem 6.34. Use the fact that $\displaystyle\int_0^1 y^x dy = \frac{1}{x+1}$ for $x > -1$ to deduce that

$$\int_0^1 y^x (-\log y)^m dy = \frac{m!}{(x+1)^{m+1}} \qquad \forall\, x > -1\,.$$

Problem 6.35. Verify that $\displaystyle\frac{d^n}{dt^n}\Gamma(t) = \int_0^\infty x^{t-1} (\log x)^n e^{-x} dx$.

Problem 6.36. In each of the following problems, find an explicit expression for F.

1. $F(x) = \displaystyle\int_0^1 \frac{y^x - 1}{\log y}\, dy.$

2. $F(x) = \displaystyle\int_0^\infty \frac{e^{-y}\big(1-\cos(xy)\big)}{y}\, dy.$

3. $F(x) = \displaystyle\int_0^\infty \frac{e^{-y}\big(1-\cos(xy)\big)}{y^2}\, dy.$

4. $F(x) = \displaystyle\int_0^\infty \frac{\sin(xy)}{y(y^2+1)} dy.$

Problem 6.37. In each of the following problems, find the range of x so that $F(x)$ is well-defined, and compute $F'(x)$.

1. $F(x) = \displaystyle\int_0^1 \frac{(\log y)^2}{1+xy} dy.$
2. $F(x) = \displaystyle\int_0^1 \frac{1}{y}\sin(xy)(\log y) dy.$
3. $F(x) = \displaystyle\int_0^\infty \frac{e^{-y}}{1+xy} dy.$
4. $F(x) = \displaystyle\int_0^\infty \frac{\sin(xy)}{y(1+y^2)} dy.$

Problem 6.38. Given the ***Legendre polynomial*** $P_n(x)$ defined by

$$P_n(x) = \frac{(-1)^n}{n!\sqrt{\pi}} \int_{-\infty}^\infty e^{-y^2(1-x^2)} \frac{d^n}{dx^n}(e^{-x^2y^2}) dy\,,$$

show that

$$x\frac{d}{dx}P_n(x) - \frac{d}{dx}P_{n-1}(x) = nP_n(x)\,.$$

Problem 6.39. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x,y) = \left\{ \begin{array}{cl} \dfrac{e^{-xy}\sin y}{y} & \text{if } y \neq 0\,, \\ 1 & \text{if } y = 0\,. \end{array} \right. .$$

Complete the following.

1. Show that $f_x(x,y)$ is continuous everywhere, and show that $f(x,\cdot)$ is integrable on $[0,\infty)$ for all $x > 0$.

2. Define $F(x) = \displaystyle\int_0^\infty f(x,y)\, dy$ for $x > 0$. Show that $F'(x) = -\dfrac{1}{x^2+1}$.

3. Show that $F(x) = \dfrac{\pi}{2} - \tan^{-1} x$ if $x > 0$, and conclude that
$$\int_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}\,.$$

Problem 6.40. Let $\{a_n\}_{n=1}^\infty \subseteq \mathbb{R}$ ba a sequence. A series $\sum\limits_{n=1}^\infty b_n$ is said to be a rearrangement of the series $\sum\limits_{n=1}^\infty a_n$ if there exists a rearrangement π of $\mathbb{N}$; that is, $\pi : \mathbb{N} \to \mathbb{N}$ is bijective, such that $b_n = a_{\pi(n)}$. Show that if $\sum\limits_{n=1}^\infty a_n$ converges absolutely, then any rearrangement of the series $\sum\limits_{n=1}^\infty a_n$ converges and has the value $\sum\limits_{n=1}^\infty a_n$.

Problem 6.41. A function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ (or $\mathbb{C}$) is called a ***double sequence***. Write $f(m,n) = a_{mn}$. A double sequence $\{a_{mn}\}_{m,n=1}^\infty$ is said to converge to a, denoted by $\lim\limits_{m,n\to\infty} a_{mn} = a$, if for every $\varepsilon > 0$, there exists $N > 0$ such that
$$|a_{mn} - a| < \varepsilon \quad \text{whenever} \quad m, n \geqslant N\,.$$
A ***double series*** is the sum of a double sequence given by $S_{pq} = \sum\limits_{m=1}^p \sum\limits_{n=1}^q a_{mn}$. The double series is said to converge to the sum S if $\lim\limits_{p,q\to\infty} S_{pq} = S$, and in this case S is usually written as $\sum\limits_{m,n=1}^\infty a_{mn}$. The double series $\sum\limits_{m=1}^p \sum\limits_{n=1}^q a_{mn}$ is said to converge absolutely if $\sum\limits_{m=1}^p \sum\limits_{n=1}^q |a_{mn}|$ converges.

Complete the following.

1. Assume that $\lim\limits_{m,n\to\infty} a_{mn} = a$. For each fixed m, assume that the limit $\lim\limits_{n\to\infty} a_{mn}$ exists. Show that $\lim\limits_{m\to\infty} \big(\lim\limits_{n\to\infty} a_{mn}\big)$ also exists and has the value a.

2. If the double series $\sum\limits_{m=1}^p \sum\limits_{n=1}^q a_{mn}$ converges absolutely, then
$$\sum_{m,n=1}^\infty a_{mn} = \sum_{m=1}^\infty \Big(\sum_{n=1}^\infty a_{mn}\Big) = \sum_{n=1}^\infty \Big(\sum_{m=1}^\infty a_{mn}\Big)\,.$$

§ 6.7 Chapter Review

Problem 6.42 (True or False). Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. Let I_1 and I_2 be open intervals in $\mathbb{R}$. Then $f : I_1 \to I_2$ is a diffeomorphism if and only if f is differentiable and $f'(x) \neq 0$ for all $x \in I_1$.

2. Let $f : [a,b] \to \mathbb{R}$ be a function. If f^2 is Riemann integrable, then f is Riemann integrable.

3. Let $f : [a,b] \to \mathbb{R}$ be a function. If f is Riemann integrable, then $\sqrt[3]{f}$ is Riemann integrable.

4. Let $f(x) = \sin \dfrac{1}{x}$ be defined on $(0,1]$. Then no matter how we define $f(0)$, f is always Riemann integrable on $[0,1]$.

5. Let $A \subseteq \mathbb{R}^n$ be bounded, and $f : A \to \mathbb{R}$ be Riemann integrable. If $\mathcal{P}$ be a partition of A, and $m \leqslant f(x) \leqslant M$ for all $x \in A$. Then $m\nu(A) \leqslant L(f,\mathcal{P}) \leqslant U(f,\mathcal{P}) \leqslant M\nu(A)$.

6. Let $A \subseteq \mathbb{R}^n$ be a set of measure zero. If $\bar{A}\backslash A$ is countable, then A has volume zero.

7. Let $A \subseteq \mathbb{R}^n$ be a closed rectangle and $f, g : B \to \mathbb{R}$ be Riemann integrable functions. If there exists a set $Z \subseteq A$ such that Z has measure zero and $g(x) = f(x)$ for all $x \in A\backslash Z$, then $\displaystyle\int_A f(x)\,dx = \int_A g(x)\,dx$.

8. Let $A \subseteq \mathbb{R}^n$ be a closed rectangle. Suppose that f and g are two bounded real-valued functions defined on A such that f is continuous and $g = f$ except on a set of measure zero, then f and g are both Riemann integrable on A.

9. Let $A, B \subseteq \mathbb{R}$ be bounded, and $f : A \to \mathbb{R}$ and $g : f(A) \to \mathbb{R}$ be Riemann integrable. Then $g \circ f$ is Riemann integrable on A.

Chapter 7

Uniform Convergence and the Space of Continuous Functions

本章的核心主題是「均勻收斂」(uniform convergence) 與「函數逼近」(approximation of functions)。

在前面幾章的分析課程中，我們通常研究單一函數的極限、連續性及可微性。然而在更高層次的觀點下，分析真正關心的是「函數空間」的結構：我們不僅研究一個函數，而是研究整個函數族 (family of functions)，以及它們之間的逼近關係。此時，均勻收斂成為關鍵工具，因為它確保極限過程能保留連續性等重要性質，使我們得以在函數空間中討論 closure 與稠密性 (density) 等概念。

本章的發展脈絡如下：

1. 建立均勻收斂的基本性質，並理解其與逐點收斂的差異；

2. 引入有界連續函數空間，將連續函數視為賦範空間中的元素；

3. 證明 Stone–Weierstrass 定理，說明在適當條件下，由簡單函數（例如多項式）所生成的代數，在 uniform norm 下對所有連續函數是稠密的。

Stone–Weierstrass 定理揭示了一個重要思想：

> 只要找到一族具有良好結構的函數，就可以均勻地逼近任意連續函數。

在這樣的脈絡下，我們進一步介紹兩個與現代人工智慧密切相關的表示定理：Universal Approximation Theorem 以及 Kolmogorov–Arnold Representation Theorem：

1. Universal Approximation Theorem 指出，由單一隱藏層神經網路所生成的函數族，在 uniform norm 下對連續函數是稠密的。從逼近理論的角度來看，這其實是一個

「ridge functions 的稠密性定理」，其本質與 Stone–Weierstrass 定理的精神一脈相承：我們尋找一族結構簡單的函數，並證明其 closure 等於整個連續函數空間。

2. Kolmogorov–Arnold Representation Theorem 則顯示，多變量連續函數可以表示為有限個單變量連續函數的組合。這說明「組合結構」本身具有強大的表示能力，也為後來的深度神經網路提供了理論上的啟發。

因此，本章並非在介紹機器學習技術，而是從純分析的觀點說明：

現代人工智慧中的某些核心表示結構，其數學基礎正建立在均勻收斂與函數逼近理論之上。

從多項式逼近，到 ridge 函數的線性組合，再到單變量函數的組合表示，我們將看到同一個核心問題反覆出現：

哪些簡單函數族，可以在 uniform norm 下逼近所有連續函數？

這正是本章的主線。

7.1 Pointwise Convergence and Uniform Convergence（逐點收斂與均勻收斂）

Definition 7.1. Let (M,d) and (N,ρ) be two metric spaces, $A \subseteq M$ be a set, and $f_k : A \to N$ be a function for each $k \in \mathbb{N}$. The sequence of functions $\{f_k\}_{k=1}^\infty$ is said to ***converge pointwise*** if $\big\{f_k(a)\big\}_{k=1}^\infty$ converges for all $a \in A$. In other words, $\{f_k\}_{k=1}^\infty$ converges pointwise if there exists a function $f : A \to N$ such that

$$\lim_{k\to\infty} \rho\big(f_k(a), f(a)\big) = 0 \qquad \forall\, a \in A.$$

In this case, $\{f_k\}_{k=1}^\infty$ is said to converge pointwise to f and is denoted by $f_k \to f$ p.w. or $f_k \overset{\text{p.w.}}{\to} f$.

Let $B \subseteq A$ be a subset. The sequence of functions $\{f_k\}_{k=1}^\infty$ is said to ***converge uniformly*** on B if there exists $f : B \to N$ such that

$$\lim_{k\to\infty} \sup_{x\in B} \rho\big(f_k(x), f(x)\big) = 0.$$

In this case, $\{f_k\}_{k=1}^{\infty}$ is said to converge uniformly to f on B (or converge to f uniformly on B) and is denoted by $f_k \to f$ unif. or $f_k \overset{\text{unif.}}{\to} f$. In other words, $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on B if for every $\varepsilon > 0$, there exists $N > 0$ such that

$$\rho\big(f_k(x), f(x)\big) < \varepsilon \qquad \forall\, k \geqslant N \text{ and } x \in B\,.$$

The sequence of functions $\{f_k\}_{k=1}^{\infty}$ is said to converge uniformly (to f) if $\{f_k\}_{k=1}^{\infty}$ converges uniformly (to f) on A.

Example 7.2. Let $f_k, f : [0,1] \to \mathbb{R}$ be given by

$$f_k(x) = \begin{cases} 0 & \text{if } \dfrac{1}{k} \leqslant x \leqslant 1\,, \\ -kx+1 & \text{if } 0 \leqslant x < \dfrac{1}{k}. \end{cases} \quad \text{and} \quad f(x) = \begin{cases} 0 & \text{if } x \in (0,1], \\ 1 & \text{if } x = 0. \end{cases}$$

Then $\{f_k\}_{k=1}^{\infty}$ converges pointwise to f. To see this, fix $x \in [0,1]$.

1. Case $x \neq 0$: Let $\varepsilon > 0$ be given, take $N \geqslant \dfrac{1}{x}$. If $k \geqslant N$, $x \geqslant \dfrac{1}{k}$ so that $f_k(x) = 0$; thus
$$\big|f_k(x) - f(x)\big| = \big|f_k(x) - 0\big| = |0 - 0| < \varepsilon\,.$$

2. Case $x = 0$: For any $\varepsilon > 0$, $k = 1, 2, 3, \dots$, $\big|f_k(0) - f(0)\big| = |1 - 1| = 0 < \varepsilon$.

However, $\{f_k\}_{k=1}^{\infty}$ does not converge uniformly to f on $[0,1]$ because

$$\lim_{k\to\infty} \sup_{x\in[0,1]} \big|f_k(x) - f(x)\big| = \lim_{k\to\infty} \sup_{x\in(0,1]} \big|f_k(x)\big| = 1 \neq 0\,.$$

Nevertheless, if $0 < a < 1$, then by the fact that f_k is decreasing for each $k \in \mathbb{N}$ and $\{f_k\}_{k=1}^{\infty}$ converges pointwise to f,

$$\lim_{k\to\infty} \sup_{x\in[a,1]} \big|f_k(x) - f(x)\big| = \lim_{k\to\infty} \sup_{x\in[a,1]} \big|f_k(x)\big| = \lim_{k\to\infty} \big|f_k(a)\big| = \big|f(a)\big| = 0$$

which implies that $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on $[a,1]$ for all $a \in (0,1)$.

Example 7.3. Let $f_k : [0,1] \to \mathbb{R}$ be given by

$$f_k(x) = \begin{cases} 0 & \text{if } \dfrac{2}{k} \leqslant x \leqslant 1\,, \\ -k^2x + 2k & \text{if } \dfrac{1}{k} \leqslant x < \dfrac{2}{k}\,, \\ k^2x & \text{if } 0 \leqslant x < \dfrac{1}{k}. \end{cases}$$

Then similar to the previous example, we have

1. $\{f_k\}_{k=1}^\infty$ converges pointwise to the zero function.
2. $\{f_k\}_{k=1}^\infty$ does not converge uniformly to the zero function on $[0,1]$.
3. $\{f_k\}_{k=1}^\infty$ converges uniformly to the zero function on $[a,1]$ for all $a \in (0,1)$.

Example 7.4. Let $f_k : [0,1] \to \mathbb{R}$ be given by $f_k(x) = x^k$. Then for each $a \in [0,1)$, $f_k(a) \to 0$ as $k \to \infty$, while if $a = 1$, $f_k(a) = 1$ for all k. Therefore, if $f(x) = \begin{cases} 0 & \text{if } x \in [0,1)\,, \\ 1 & \text{if } x = 1\,, \end{cases}$ then $f_k \to f$ p.w.. However, since

$$\sup_{x\in[0,1]} \big|f_k(x) - f(x)\big| = \sup_{x\in[0,1)} \big|f_k(x)\big| = 1\,,$$

we must have

$$\lim_{k\to\infty} \sup_{x\in[0,1]} \big|f_k(x) - f(x)\big| = 1 \neq 0\,.$$

Therefore, $\{f_k\}_{k=1}^\infty$ does not converge uniformly to f on $[0,1]$.

On the other hand, if $0 < a < 1$, then the fact that f_k is increasing for all $k \in \mathbb{N}$ implies that

$$\sup_{x\in[0,a]} \big|f_k(x) - f(x)\big| \leqslant a^k\,;$$

thus by the Sandwich lemma,

$$\lim_{k\to\infty} \sup_{x\in[0,a]} \big|f_k(x) - f(x)\big| = 0\,.$$

Therefore, $\{f_k\}_{k=1}^\infty$ converges to uniformly f on $[0,a]$ if $0 < a < 1$.

Example 7.5. Let $f_k : \mathbb{R} \to \mathbb{R}$ be given by $f_k(x) = \dfrac{\sin x}{k}$. Then for each $x \in \mathbb{R}$, $|f_k(x)| \leqslant \dfrac{1}{k}$ which converges to 0 as $k \to \infty$. By the Sandwich lemma,

$$\lim_{k\to\infty} \big|f_k(x)\big| = 0 \qquad \forall\, x \in \mathbb{R}\,.$$

Therefore, $f_k \to 0$ p.w.. Moreover, since $\sup\limits_{x\in\mathbb{R}} \big|f_k(x)\big| \leqslant \dfrac{1}{k}$, $\lim\limits_{k\to\infty} \sup\limits_{x\in\mathbb{R}} \big|f_k(x)\big| = 0$. Therefore, $\{f_k\}_{k=1}^\infty$ converges uniformly to 0 on $\mathbb{R}$.

Proposition 7.6. *Let (M,d) and (N,ρ) be two metric spaces, $A \subseteq M$ be a set, and $f_k, f : A \to N$ be functions for $k = 1, 2, \cdots$. If $\{f_k\}_{k=1}^\infty$ converges uniformly to f on A, then $\{f_k\}_{k=1}^\infty$ converges pointwise to f.*

Proof. For each $a \in A$, $\rho\big(f_k(a), f(a)\big) \leqslant \sup_{x\in A} \rho\big(f_k(x), f(x)\big)$; thus the Sandwich lemma shows that

$$\lim_{k\to\infty} \rho\big(f_k(a), f(a)\big) = 0$$

since $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on A. □

Proposition 7.7 (Cauchy criterion for uniform convergence)**.** *Let (M, d) and (N, ρ) be two metric spaces, $A \subseteq M$ be a set, and $f_k : A \to N$ be a sequence of functions. Suppose that (N, ρ) is complete. Then $\{f_k\}_{k=1}^{\infty}$ converges uniformly on $B \subseteq A$ if and only if for every $\varepsilon > 0$, there exists $N > 0$ such that*

$$\rho\big(f_k(x), f_\ell(x)\big) < \varepsilon \qquad \forall\, k, \ell \geqslant N \text{ and } x \in B\,.$$

Proof. "⇒" Suppose that $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on B. Let $\varepsilon > 0$ be given. Then there exists $N > 0$ such that

$$\rho\big(f_k(x), f(x)\big) < \frac{\varepsilon}{2} \qquad \forall\, k \geqslant N \text{ and } x \in B\,.$$

Then if $k, \ell \geqslant N$ and $x \in B$,

$$\rho\big(f_k(x), f_\ell(x)\big) \leqslant \rho\big(f_k(x), f(x)\big) + \rho\big(f(x), f_\ell(x)\big) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

"⇐" Let $b \in B$. By assumption, $\big\{f_k(b)\big\}_{k=1}^{\infty}$ is a Cauchy sequence in (N, ρ); thus is convergent by the completeness of (N, ρ). Therefore, we establish a map $f : B \to N$ defined by $f(b) = \lim\limits_{k\to\infty} f_k(b)$. We claim that $\{f_k\}_{k=1}^{\infty}$ convergence uniformly to f on B.

Let $\varepsilon > 0$ be given. Then there exists $N > 0$ such that

$$\rho\big(f_k(x), f_\ell(x)\big) < \frac{\varepsilon}{2} \qquad \forall\, k, \ell \geqslant N \text{ and } x \in B\,.$$

Moreover, for each $x \in B$ there exists $N_x > 0$ such that

$$\rho\big(f_\ell(x), f(x)\big) < \frac{\varepsilon}{2} \qquad \forall\, \ell \geqslant N_x\,.$$

Then for all $k \geqslant N$ and $x \in B$,

$$\rho\big(f_k(x), f(x)\big) \leqslant \rho\big(f_k(x), f_\ell(x)\big) + \rho\big(f_\ell(x), f(x)\big) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

in which we choose $\ell \geqslant \max\{N, N_x\}$ to conclude the inequality. □

Theorem 7.8. *Let (M,d) and (N,ρ) be two metric spaces, $A \subseteq M$ be a set, and $f_k : A \to N$ be a sequence of continuous functions converging to $f : A \to N$ uniformly on A. Then f is continuous on A; that is,*

$$\lim_{x\to a} f(x) = \lim_{x\to a}\lim_{k\to\infty} f_k(x) = \lim_{k\to\infty}\lim_{x\to a} f_k(x) = f(a)\,.$$

Proof. Let $a \in A$ and $\varepsilon > 0$ be given. Since $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on A, there exists $N > 0$ such that

$$\rho\big(f_k(x), f(x)\big) < \frac{\varepsilon}{3} \qquad \forall\, k \geqslant N \text{ and } x \in A\,.$$

By the continuity of f_N, there exists $\delta > 0$ such that

$$\rho\big(f_N(x), f_N(a)\big) < \frac{\varepsilon}{3} \quad \text{whenever} \quad x \in B_M(a,\delta) \cap A\,.$$

Therefore, if $x \in B_M(a,\delta) \cap A$, by the triangle inequality

$$\begin{aligned}\rho\big(f(x), f(a)\big) &\leqslant \rho\big(f(x), f_N(x)\big) + \rho\big(f_N(x), f_N(a)\big) + \rho\big(f_N(a), f(a)\big)\\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon\,;\end{aligned}$$

thus f is continuous at a. □

Example 7.9. Let $f_k : [0,2] \to \mathbb{R}$ be given by $f_k(x) = \dfrac{x^k}{1+x^k}$. Then

1. For each $a \in [0,1)$, $f_k(a) \to 0$ as $k \to \infty$;
2. For each $a \in (1,2]$, $f_k(a) \to 1$ as $k \to \infty$;
3. If $a = 1$, then $f_k(a) = \dfrac{1}{2}$ for all k.

Let $f(x) = \begin{cases} 0 & \text{if } x \in [0,1)\,, \\ \dfrac{1}{2} & \text{if } x = 1\,, \\ 1 & \text{if } x \in (1,2]\,. \end{cases}$ Then $\{f_k\}_{k=1}^{\infty}$ converges pointwise to f. However, $\{f_k\}_{k=1}^{\infty}$ does not converge uniformly to f on $[0,2]$ since f_k are continuous functions for all $k \in \mathbb{N}$ but f is not.

Remark 7.10. The uniform limit of sequence of continuous function might not be uniformly continuous. For example, let $A = (0,1)$ and $f_k(x) = \dfrac{1}{x}$ for all $k \in \mathbb{N}$. Then $\{f_k\}_{k=1}^{\infty}$ converges uniformly to $f(x) = \dfrac{1}{x}$, but the limit function is not uniformly continuous on A.

Theorem 7.11. *Let $I \subseteq \mathbb{R}$ be a finite interval, $f_k : I \to \mathbb{R}$ be a sequence of differentiable functions, and $g : I \to \mathbb{R}$ be a function. Suppose that $\{f_k(a)\}_{k=1}^{\infty}$ converges for some $a \in I$, and $\{f_k'\}_{k=1}^{\infty}$ converges uniformly to g on I. Then*

1. *$\{f_k\}_{k=1}^{\infty}$ converges uniformly to some function f on I.*

2. *The limit function f is differentiable on I, and $f'(x) = g(x)$ for all $x \in I$; that is,*

$$\lim_{k\to\infty} f_k'(x) = \lim_{k\to\infty} \frac{d}{dx} f_k(x) = \frac{d}{dx} \lim_{k\to\infty} f_k(x) = f'(x).$$

Proof. 1. Let $\varepsilon > 0$ be given. Since $\{f_k(a)\}_{k=1}^{\infty}$ converges to $f(a)$, $\{f_k(a)\}_{k=1}^{\infty}$ is a Cauchy sequence. Therefore, there exists $N_1 > 0$ such that

$$\big|f_k(a) - f_\ell(a)\big| < \frac{\varepsilon}{2} \qquad \forall\, k, \ell \geqslant N_1 .$$

By the uniform convergence of $\{f_k'\}_{k=1}^{\infty}$ on I and Proposition 7.7, there exists $N_2 > 0$ such that

$$\big|f_k'(x) - f_\ell'(x)\big| < \frac{\varepsilon}{2|I|} \qquad \forall\, k, \ell \geqslant N_2 \text{ and } x \in I\,,$$

where $|I|$ is the length of the interval.

Let $N = \max\{N_1, N_2\}$. By the mean value theorem, for all $k, \ell \geqslant N$ and $x \in I$, there exists ξ in between x and a such that

$$\big|f_k(x) - f_\ell(x) - f_k(a) + f_\ell(a)\big| = |f_k'(\xi) - f_\ell'(\xi)||x - a| \leqslant \frac{\varepsilon|x-a|}{2|I|} \leqslant \frac{\varepsilon}{2};$$

thus for all $k, \ell \geqslant N$ and $x \in I$,

$$\big|f_k(x) - f_\ell(x)\big| \leqslant \big|f_k(a) - f_\ell(a)\big| + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore, Proposition 7.7 implies that $\{f_k\}_{k=1}^{\infty}$ converges uniformly on I.

2. Suppose that the uniform limit of $\{f_k\}_{k=1}^{\infty}$ is f. For any given point $c \in I$, define

$$\phi_k(x) = \begin{cases} \dfrac{f_k(x) - f_k(c)}{x - c} & \text{if } x \in I,\ x \neq c, \\ f_k'(c) & \text{if } x = c, \end{cases} \quad \text{and} \quad \phi(x) = \begin{cases} \dfrac{f(x) - f(c)}{x - c} & \text{if } x \in I,\ x \neq c, \\ g(c) & \text{if } x = c. \end{cases}$$

Then ϕ_k is continuous on I for all $k \in \mathbb{N}$, and $\{\phi_k\}_{k=1}^{\infty}$ converges pointwise to ϕ.

Claim: $\{\phi_k\}_{k=1}^{\infty}$ converges uniformly to ϕ on I.

Proof of claim: Let $\varepsilon > 0$ be given. Since $\{f_k'\}_{k=1}^{\infty}$ converges uniformly on I, there exists $N > 0$ such that

$$\sup_{s\in I}\big|f_k'(s) - f_\ell'(s)\big| < \varepsilon \qquad \forall\, k, \ell \geqslant N\,.$$

Since

$$\big|\phi_k(x) - \phi_\ell(x)\big| = \begin{cases} \dfrac{\big|f_k(x) - f_k(c) - f_\ell(x) + f_\ell(c)\big|}{|x-c|} & \text{if } x \neq c,\ x \in I\,, \\ \big|f_k'(c) - f_\ell'(c)\big| & \text{if } x = c\,, \end{cases}$$

by the mean value theorem we obtain that

$$\big|\phi_k(x) - \phi_\ell(x)\big| \leqslant \sup_{s\in I}\big|f_k'(s) - f_\ell'(s)\big| < \varepsilon \qquad \forall\, k, \ell \geqslant N \text{ and } x \in I\,.$$

Therefore, the Cauchy criterion shows that $\{\phi_k\}_{k=1}^{\infty}$ converges uniformly to ϕ on I, and Theorem 7.8 further shows that ϕ is continuous on I; thus

$$f'(c) = \lim_{x\to c}\phi(x) = \phi(c) = g(c)\,. \qquad \square$$

Example 7.12. Assume that $f_k : I \to \mathbb{R}$ is differentiable for all $k \in \mathbb{N}$, and $\{f_k'\}_{k=1}^{\infty}$ converges uniformly to g on I. Then $\{f_k\}_{k=1}^{\infty}$ might **NOT** converge. For example, consider $f_k(x) = k$. Then $f_k' \equiv 0$ but $\{f_k\}_{k=1}^{\infty}$ does not converge.

Example 7.13. For each $k \in \mathbb{N}$, let $f_k : [0,1] \to \mathbb{R}$ be defined by $f_k(x) = \dfrac{\sin(k^2x)}{k}$. Then $\{f_k\}_{k=1}^{\infty}$ converges uniformly to the zero function on $[0,1]$ since

$$\sup_{x\in[0,1]}\big|f_k(x) - 0\big| = \sup_{x\in[0,1]}\Big|\frac{\sin(k^2x)}{k}\Big| \leqslant \frac{1}{k} \Rightarrow \lim_{k\to\infty}\sup_{x\in[0,1]}\big|f_k(x) - 0\big| = 0\,.$$

However, note that $f_k'(x) = k\cos(k^2x)$ so that $\{f_k'(0) = k$ which diverges to ∞ (so that $\{f_k'\}_{k=1}^{\infty}$ does not even converge pointwise). Therefore, even if a differentiable sequence $\{f_k\}_{k=1}^{\infty}$ converges uniformly, it does not implies that $\{f_k'\}_{k=1}^{\infty}$ converges (pointwise).

Example 7.14. Assume that $f_k : I \to \mathbb{R}$ is differentiable for all $k \in \mathbb{N}$, and $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on I. Then f might **NOT** be differentiable. In fact, there are differentiable functions $f_k : [a,b] \to \mathbb{R}$ such that f_k converges uniformly to f on $[a,b]$ but f is not differentiable. For example, consider

$$f_k(x) = \begin{cases} \dfrac{k}{2}x^2 & \text{if } |x| \leqslant \dfrac{1}{k}, \\ |x| - \dfrac{1}{2k} & \text{if } \dfrac{1}{k} \leqslant |x| \leqslant 1. \end{cases}$$

Observe that $f_k(-x) = f_k(x)$, so it suffices to consider $x \geqslant 0$.

1. Let $f(x) = |x|$. Then $f_k \to f$ uniformly:

$$\begin{aligned} &\sup_{x\in[-1,1]} \big|f_k(x) - f(x)\big| \\ &\qquad = \sup_{x\in[0,1]} \big|f_k(x) - x\big| = \max\Big\{\sup_{x\in[0,\frac{1}{k}]} \big|f_k(x) - x\big|, \sup_{x\in[\frac{1}{k},1]} \big|f_k(x) - x\big|\Big\} \\ &\qquad = \max\Big\{\sup_{x\in[0,\frac{1}{k}]} \Big|\frac{kx^2}{2} - x\Big|, \sup_{x\in[\frac{1}{k},1]} \Big|x - \frac{1}{2k} - x\Big|\Big\} \\ &\qquad \leqslant \sup_{x\in[0,\frac{1}{k}]} \Big|\frac{kx^2}{2}\Big| + |x| \leqslant \frac{k}{2}\Big(\frac{1}{k}\Big)^2 + \frac{1}{k} = \frac{3}{2k} \to 0 \text{ as } k \to \infty. \end{aligned}$$

2. To see if f_k are differentiable, it suffices to show $f_k'(\frac{1}{k})$ exists.

$$f_k'\Big(\frac{1}{k}\Big) = \lim_{h\to 0} \frac{f_k(\frac{1}{k}+h) - f_k(\frac{1}{k})}{h} = \lim_{h\to 0} \frac{1}{h} \begin{cases} \big(\dfrac{1}{k} + h\big) - \dfrac{1}{2k} - \dfrac{1}{2k} & \text{if } h > 0 \\ \dfrac{k}{2}\Big(\dfrac{1}{k} + h\Big)^2 - \dfrac{1}{2k} & \text{if } h < 0 \end{cases}$$

$$= \lim_{h\to 0} \frac{1}{h} \begin{cases} h & \text{if } h > 0 \\ h + \dfrac{k}{2}h^2 & \text{if } h < 0 \end{cases} = 1.$$

Example 7.15. Assume that $f_k : [-1,1] \to \mathbb{R}$ be given by

$$f_k(x) = \begin{cases} 0 & \text{if } x \in [-1,0], \\ \dfrac{k^2}{2}x^2 & \text{if } x \in \big(0, \dfrac{1}{k}\big], \\ 1 - \dfrac{k^2}{2}\big(x - \dfrac{2}{k}\big)^2 & \text{if } x \in \big(\dfrac{1}{k}, \dfrac{2}{k}\big], \\ 1 & \text{if } x \in \big(\dfrac{2}{k}, 1\big]. \end{cases}$$

Then $f_k'(x) = \begin{cases} 0 & \text{if } x \in [-1,0], \\ k^2 x & \text{if } x \in \big(0, \frac{1}{k}\big], \\ -k^2\big(x - \frac{2}{k}\big) & \text{if } x \in \big(\frac{1}{k}, \frac{2}{k}\big], \\ 0 & \text{if } x \in \big(\frac{2}{k}, 1\big], \end{cases}$ and $\{f_k'\}_{k=1}^\infty$ converges pointwise to 0 but not uniformly on $[-1,1]$. We note that $\{f_k\}_{k=1}^\infty$ converges to a discontinuous function

$$f(x) = \begin{cases} 0 & \text{if } x \in [-1,0], \\ 1 & \text{if } x \in (0,1], \end{cases}$$

so the convergence of $\{f_k\}_{k=1}^\infty$ cannot be uniform on $[-1,1]$.

Example 7.16. There are differentiable functions $f_k : [a,b] \to \mathbb{R}$ such that f_k converges uniformly to f on $[a,b]$ but $\lim\limits_{k\to\infty} f_k' \neq (\lim\limits_{k\to\infty} f_k)'$. For example, take $f_k(x) = \dfrac{x}{1+k^2x^2}$ on $[-1,1]$. Then $f_k'(x) = \dfrac{1-k^2x^2}{(1+k^2x^2)^2}$.

1. Since $\lim\limits_{k\to\infty} \sup\limits_{x\in[-1,1]} \Big|\dfrac{x}{1+k^2x^2} - 0\Big| = \lim\limits_{k\to\infty} \dfrac{1}{2k} = 0$, f_k converges uniformly to 0 on $[-1,1]$.
2. $\big(\lim\limits_{k\to\infty} f_k(x)\big)' = 0' = 0$.
3. $\lim\limits_{k\to\infty} f_k'(x) = \lim\limits_{k\to\infty} \dfrac{1-k^2x^2}{(1+k^2x^2)^2} = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x \neq 0, |x| < 1. \end{cases}$ Note that f_k' does not converge uniformly.

Theorem 7.17. *Let $f_k : A \to \mathbb{R}$ be a sequence of Riemann integrable functions which converges uniformly to f on A. Then f is Riemann integrable on A, and*

$$\lim_{k\to\infty} \int_A f_k(x)\,dx = \int_A \lim_{k\to\infty} f_k(x)\,dx = \int_A f(x)\,dx. \tag{7.1.1}$$

Proof. Let R be a rectangle such that $A \subseteq R$ and $\nu(R) > 0$, and $\varepsilon > 0$ be given. Since $\{f_k\}_{k=1}^\infty$ converges uniformly to f on A, there exists $N > 0$ such that

$$\big|f_k(x) - f(x)\big| < \frac{\varepsilon}{4\nu(R)} \qquad \forall\, k \geqslant N \text{ and } x \in A. \tag{7.1.2}$$

Since f_N is Riemann integrable on A, by Riemann's condition there exists a partition $\mathcal{P}$ of A such that

$$U(f_N, \mathcal{P}) - L(f_N, \mathcal{P}) < \frac{\varepsilon}{2}.$$

By Proposition 6.8, we find that

$$\begin{aligned} U(f,\mathcal{P}) - L(f,\mathcal{P}) &= U(f - f_{N} + f_{N}, \mathcal{P}) - L(f - f_{N} + f_{N}, \mathcal{P}) \\ &\leqslant U(f - f_{N}, \mathcal{P}) + U(f_{N}, \mathcal{P}) - L(f - f_{N}, \mathcal{P}) - L(f_{N}, \mathcal{P}) \\ &\leqslant \frac{\varepsilon}{4\nu(R)}\nu(R) + \frac{\varepsilon}{4\nu(R)}\nu(R) + U(f_{N}, \mathcal{P}) - L(f_{N}, \mathcal{P}) \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{2} = \varepsilon; \end{aligned}$$

thus by Riemann's condition f is Riemann integrable on A.

Now, if $k \geqslant N$, (7.1.2) implies that

$$\begin{aligned} \Big|\int_A f_k(x)dx - \int_A f(x)dx\Big| &= \Big|\int_A \big(f_k(x) - f(x)\big)dx\Big| \leqslant \int_A \big|f_k(x) - f(x)\big|dx \\ &\leqslant \frac{\varepsilon}{4\nu(R)}\nu(R) = \frac{\varepsilon}{4} < \varepsilon \end{aligned}$$

which shows (7.1.1). □

Example 7.18. In this example we provide a sequence of integrable functions converges pointwise to a limit function which is not integrable. Let $\{q_k\}_{k=1}^{\infty}$ be the rational numbers in $[0,1]$, and

$$f_k(x) = \begin{cases} 0 & \text{if } x \in \{q_1, q_2, \cdots, q_k\}, \\ 1 & \text{otherwise}. \end{cases}$$

Then f_k converges pointwise to the Dirichlet function

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q} \cap [0,1], \\ 1 & \text{if } x \in [0,1]\backslash\mathbb{Q}. \end{cases}$$

It is well-known that the Dirichlet function is not integrable. However, $\{f_k\}_{k=1}^{\infty}$ does not converge uniformly to f since f_k are Riemann integrable on $[0,1]$ for all $k \in \mathbb{N}$ but f is not.

7.2 Series of Functions and The Weierstrass M-Test

Definition 7.19. Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed space, $A \subseteq M$ be a subset, and $g_k, g : A \to \mathcal{V}$ be functions. We say that the series $\sum_{k=1}^{\infty} g_k$ converges pointwise if the sequence of partial sum $\{s_n\}_{n=1}^{\infty}$ given by

$$s_n = \sum_{k=1}^{n} g_k$$

converges pointwise. We use $\sum\limits_{k=1}^{\infty} g_k = g$ p.w. to denote that the series $\sum\limits_{k=1}^{\infty} g_k$ converges pointwise to g. The series $\sum\limits_{k=1}^{\infty} g_k$ is said to converge uniformly on $B \subseteq A$ if $\{s_n\}_{n=1}^{\infty}$ converges uniformly on B.

Example 7.20. Consider the geometric series $\sum\limits_{k=0}^{\infty} x^k$. The partial sum s_n is given by

$$s_n(x) = \begin{cases} \dfrac{1-x^{n+1}}{1-x} & \text{if } x \neq 1\,, \\ n+1 & \text{if } x = 1\,. \end{cases}$$

Then

1. $\sum\limits_{k=0}^{\infty} x^k$ converges pointwise to $g(x) = \dfrac{1}{1-x}$ in $(-1,1)$.

2. $\sum\limits_{k=0}^{\infty} x^k$ does not converge pointwise in $(-\infty,-1] \cup [1,\infty)$.

3. $\sum\limits_{k=0}^{\infty} x^k$ converges uniformly on $(-a,a)$ if $0 < a < 1$ since
$$\sup_{x\in(-a,a)} |s_n(x) - g(x)| = \sup_{x\in(-a,a)} \frac{|x|^{n+1}}{1-x} \leqslant \frac{|a|^{n+1}}{1-a} \to 0 \text{ as } n \to \infty.$$

4. $\sum\limits_{k=0}^{\infty} x^k$ does not converge uniformly on $(-1,1)$ since $\sup\limits_{x\in(-1,1)} |s_n(x) - g(x)| = \infty$.

The following two corollaries are direct consequences of Proposition 7.7 and Theorem 7.8.

Corollary 7.21. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a complete normed vector space, $A \subseteq M$ be a subset, and $g_k : A \to \mathcal{V}$ be functions. Then $\sum\limits_{k=1}^{\infty} g_k$ converges uniformly on A if and only if*

$$\forall\, \varepsilon > 0, \exists\, N > 0 \ni \Big\| \sum_{k=m+1}^{n} g_k(x) \Big\| < \varepsilon \qquad \forall\, n > m \geqslant N \text{ and } x \in A.$$

Corollary 7.22. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, $A \subseteq M$ be a subset, and $g_k, g : A \to \mathcal{V}$ be functions. If $g_k : A \to \mathcal{V}$ is continuous for all $k \in \mathbb{N}$ and $\sum\limits_{k=1}^{\infty} g_k(x)$ converges to g uniformly on A, then g is continuous.*

Theorem 7.23. *Let $f:(a,b)\to\mathbb{R}$ be an infinitely differentiable functions; that is, $f^{(k)}(x)$ exists for all $k\in\mathbb{N}$ and $x\in(a,b)$. Let $c\in(a,b)$ and suppose that for some $0<h<\infty$, $\big|f^{(k)}(x)\big|\leqslant M$ for all $x\in(c-h,c+h)\subseteq(a,b)$ and $k\in\mathbb{N}$. Then*

$$f(x)=\sum_{k=0}^{\infty}\frac{f^{(k)}(c)}{k!}(x-c)^k \qquad \forall\, x\in(c-h,c+h)$$

and the convergence is uniform.

Proof. First, we claim that

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(c)}{k!}(x-c)^k+(-1)^n\int_c^x\frac{(y-x)^n}{n!}f^{(n+1)}(y)dy \qquad \forall\, x\in(a,b). \tag{7.2.1}$$

By the fundamental theorem of Calculus it is clear that (7.2.1) holds for $n=0$. Suppose that (7.2.1) holds for $n=m$. Then

$$\begin{aligned} f(x)&=\sum_{k=0}^{m}\frac{f^{(k)}(c)}{k!}(x-c)^k+(-1)^m\Big[\frac{(y-x)^{m+1}}{(m+1)!}f^{(m+1)}(y)\Big|_{y=c}^{y=x}-\int_c^x\frac{(y-x)^{m+1}}{(m+1)!}f^{(m+2)}(y)dy\Big]\\ &=\sum_{k=0}^{m+1}\frac{f^{(k)}(c)}{k!}(x-c)^k+(-1)^{m+1}\int_c^x\frac{(y-x)^{m+1}}{(m+1)!}f^{(m+2)}(y)dy \end{aligned}$$

which implies that (7.2.1) also holds for $n=m+1$. By induction (7.2.1) holds for all $n\in\mathbb{N}$.

Define $s_n(x)=\sum\limits_{k=0}^{n}\dfrac{f^{(k)}(c)}{k!}(x-c)^k$. Our goal is to show that $\{s_n\}_{n=1}^{\infty}$ converges uniformly to f on $(c-h,c+h)$. Nevertheless, note that if $x\in(c-h,c+h)$,

$$\big|s_n(x)-f(x)\big|\leqslant\Big|\int_c^x\frac{h^n}{n!}M\,dy\Big|\leqslant\frac{h^{n+1}}{n!}M.$$

By the fact that $\lim\limits_{n\to\infty}\dfrac{h^{n+1}}{n!}M=0$, we conclude that

$$\lim_{n\to\infty}\sup_{x\in(c-h,c+h)}\big|s_n(x)-f(x)\big|=0. \qquad \square$$

Remark 7.24. Assume the conditions in Theorem 7.23. Then applying Theorem 7.23 to the function $g=f'$ shows that

$$\frac{d}{dx}\sum_{k=0}^{\infty}\frac{f^{(k)}(c)}{k!}(x-c)^k=\sum_{k=0}^{\infty}\frac{d}{dx}\Big[\frac{f^{(k)}(c)}{k!}(x-c)^k\Big] \qquad \forall\, x\in(c-h,c+h).$$

This is a special case of Corollary 7.38.

Example 7.25. The series $\sum_{k=0}^{\infty}(-1)^k\dfrac{x^{2k+1}}{(2k+1)!}$ converges to $\sin x$ uniformly on any bounded subset of $\mathbb{R}$.

Theorem 7.26 (Weierstrass M-test)**.** *Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a complete normed vector space, $A\subseteq M$ be a subset, and $g_k : A\to\mathcal{V}$ be a sequence of functions. Suppose that there exists $M_k>0$ such that $\sup_{x\in A}\|g_k(x)\|\leqslant M_k$ for all $k\in\mathbb{N}$ and $\sum_{k=1}^{\infty}M_k$ converges. Then $\sum_{k=1}^{\infty}g_k$ converges uniformly and absolutely (that is, $\sum_{k=1}^{\infty}\|g_k\|$ converges uniformly) on A.*

Proof. We show that the partial sum $s_n=\sum_{k=1}^{n}g_k$ satisfies the Cauchy criterion. Let $\varepsilon>0$ be given. Since $\sum_{k=1}^{\infty}M_k$ converges (which means $\sum_{k=1}^{n}M_k$ converges as $n\to\infty$), the Cauchy criterion for the convergence of series of vectors (Theorem 2.65) implies that there exists $N>0$ such that

$$\sum_{k=m+1}^{n}M_k=\Big|\sum_{k=m+1}^{n}M_k\Big|<\varepsilon\quad\forall\,n>m\geqslant N\,.$$

Therefore,

$$\Big\|\sum_{k=m+1}^{n}g_k(x)\Big\|\leqslant\sum_{k=m+1}^{n}\big\|g_k(x)\big\|\leqslant\sum_{k=m+1}^{n}M_k<\varepsilon\quad\forall\,n>m\geqslant N\text{ and }x\in A$$

and the desired result follows from the Cauchy criterion for the uniform convergence of series of functions (Corollary 7.21). □

Theorem 7.8 and 7.26 together imply the following

Corollary 7.27. *Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a complete normed vector space, $A\subseteq M$ be a subset, and $g_k : A\to\mathcal{V}$ be a sequence of continuous functions. Suppose that there exists $M_k>0$ such that $\sup_{x\in A}\|g_k(x)\|\leqslant M_k$ for all $k\in\mathbb{N}$ and $\sum_{k=1}^{\infty}M_k$ converges. Then $\sum_{k=1}^{\infty}g_k$ is continuous on A.*

Example 7.28. Consider the series $f(x)=\sum_{k=0}^{\infty}\Big(\dfrac{x^k}{k!}\Big)^2$. For all $x\in[-R,R]$, $\Big(\dfrac{x^k}{k!}\Big)^2\leqslant\dfrac{R^{2k}}{(k!)^2}$. Moreover,

$$\limsup_{k\to\infty}\frac{R^{2(k+1)}}{((k+1)!)^2}\Big/\frac{R^{2k}}{(k!)^2}=\limsup_{k\to\infty}\frac{R^2}{(k+1)^2}=0;$$

thus the ratio test and the Weierstrass M-test imply that the series $\sum_{k=0}^{\infty}\Big(\dfrac{x^k}{k!}\Big)^2$ converges uniformly on $[-R, R]$. Theorem 7.8 then shows that f is continuous on $[-R, R]$. Since R is arbitrary, we find that f is continuous on $\mathbb{R}$.

Example 7.29. Let $\{a_k\}_{k=0}^{\infty}$ be a bounded sequence. Then $\sum\limits_{k=0}^{\infty} \dfrac{a_k}{k!} x^k$ converges to a continuous function.

Example 7.30. Consider the function $f(x) = \dfrac{\pi}{2} - \dfrac{4}{\pi} \sum\limits_{k=0}^{\infty} \dfrac{\cos(2k+1)x}{(2k+1)^2}$. We can in fact show (later) that $f(x) = |x|$ for all $x \in [-\pi, \pi]$, and by the Weierstrass M-test it is easy to see that the convergence is uniform on $\mathbb{R}$.

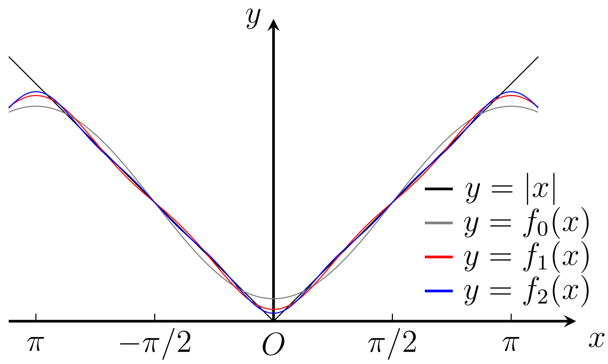


Figure 7.1: The graph of some partial sums

7.3 Integration and Differentiation of Series

The following two theorems are direct consequences of Theorem 7.11 and 7.17.

Theorem 7.31. *Let $g_k : [a, b] \to \mathbb{R}$ be a sequence of Riemann integrable functions. If $\sum\limits_{k=1}^{\infty} g_k$ converges uniformly on $[a, b]$, then*

$$\int_a^b \sum_{k=1}^{\infty} g_k(x)dx = \sum_{k=1}^{\infty} \int_a^b g_k(x)dx .$$

Theorem 7.32. *Let $g_k : (a, b) \to \mathbb{R}$ be a sequence of differentiable functions. Suppose that $\sum\limits_{k=1}^{\infty} g_k$ converges for some $c \in (a, b)$, and $\sum\limits_{k=1}^{\infty} g_k'$ converges uniformly on (a, b). Then*

$$\sum_{k=1}^{\infty} g_k'(x) = \frac{d}{dx} \sum_{k=1}^{\infty} g_k(x) .$$

Definition 7.33. A series is called a ***power series about*** c or ***centered at*** c if it is of the form $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ for some sequence $\{a_k\}_{k=0}^{\infty} \subseteq \mathbb{R}$ (or $\mathbb{C}$) and $c \in \mathbb{R}$ (or $\mathbb{C}$).

Proposition 7.34. *If a power series centered at c is convergent at some point $b \neq c$, then the power series converges pointwise on $B(c, |b-c|)$, and converges uniformly on any compact subsets of $B(c, |b-c|)$.*

Proof. Since the series $\sum\limits_{k=0}^{\infty} a_k(b-c)^k$ converges, $|a_k||b-c|^k \to 0$ as $k \to \infty$; thus there exists $M > 0$ such that $|a_k||b-c|^k \leqslant M$ for all k.

1. $x \in B(c, |b-c|)$, the series $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ converges absolutely since
$$\sum_{k=0}^{\infty} |a_k(x-c)^k| \leqslant \sum_{k=0}^{\infty} |a_k||x-c|^k = \sum_{k=0}^{\infty} |a_k||b-c|^k \frac{|x-c|^k}{|b-c|^k} \leqslant M \sum_{k=0}^{\infty} \Big(\frac{|x-c|}{|b-c|}\Big)^k$$
which converges (because of the geometric series test or ratio test).

2. Let $K \subseteq B(c, |b-c|)$ be a compact set. Then
$$\operatorname{dist}(K, \partial B(c, |b-c|)) \equiv \inf\big\{|x-y| \,\big|\, x \in K, |y-c| = |b-c|\big\} > 0.$$
Define $r = \dfrac{|b-c| - \operatorname{dist}(K, \partial B(c, |b-c|))}{|b-c|}$. Then $0 \leqslant r < 1$, and $|x-c| \leqslant r|b-c|$ for all $x \in K$. Therefore, $|a_k(x-c)^k| \leqslant Mr^k$ if $x \in K$; thus the Weierstrass M-test implies that the series $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ converges uniformly on K. ◻

By the proposition above, we immediately conclude that the collection of all x at which the power series converges must be connected and symmetric; thus is a disc or a point. This observation induce the following

Definition 7.35. A non-negative number R is called the ***radius of convergence*** of the power series $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ if the series converges for all $x \in B(c, R)$ but diverges if $x \notin B[c, R]$. In other words,
$$R = \sup\Big\{r \geqslant 0 \,\Big|\, \sum_{k=0}^{\infty} a_k(x-c)^k \text{ converges in } B[c, R]\Big\}.$$
The ***interval of convergence*** or ***convergence interval*** of a power series is the collection of all x at which the power series converges.

Remark 7.36. A power series converges pointwise on its interval of convergence.

Theorem 7.37. *Let $\{a_k\}_{k=0}^{\infty} \subseteq \mathbb{C}$, $c \in \mathbb{C}$, $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ be a power series with radius of convergence $R > 0$, and $K \subseteq B(c, R)$ be a compact set. Then*

1. *The power series* $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ *converges uniformly on* K.

2. *The power series* $\sum\limits_{k=0}^{\infty} (k+1)a_{k+1}(x-c)^k$ *converges pointwise on* $B(c,R)$*, and converges uniformly on* K.

Proof. 1. It is simply a restatement of Proposition 7.34.

2. By 1, it suffices to show that the power series $\sum\limits_{k=0}^{\infty}(k+1)a_{k+1}(x-c)^k$ converges pointwise on $B(c,R)$. Clearly the series converges at $x=c$. Let $x \in B(c,R)$ and $x \neq c$. Since $|x-c|<R$, there exists $b \in B(c,R)$ such that
$$|b-c| = \frac{R+|x-c|}{2}.$$
Then if $r = \dfrac{|x-c|}{|b-c|}$, $0<r<1$ and
$$\sum_{k=0}^{\infty}(k+1)|a_{k+1}||x-c|^k \leqslant \sum_{k=0}^{\infty}(k+1)|a_{k+1}||b-c|^k\Big(\frac{|x-c|}{|b-c|}\Big)^k \leqslant M\sum_{k=0}^{\infty}(k+1)r^k$$
for some $M>0$. Note that the ratio test implies that the series $\sum\limits_{k=0}^{\infty}(k+1)r^k$ converges if $0<r<1$; thus $\sum\limits_{k=0}^{\infty}(k+1)|a_{k+1}||x-c|^k$ converges by the comparison test. □

Corollary 7.38. *Let* $\{a_k\}_{k=0}^{\infty} \subseteq \mathbb{R}$ *and* $c \in \mathbb{R}$*, and* $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ *be a power series with radius of convergence* $R>0$*. Then* $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ *is differentiable in* $(c-R,c+R)$ *and is Riemann integrable over any closed intervals* $[\alpha,\beta] \subseteq (c-R,c+R)$*. Moreover,*
$$\frac{d}{dx}\sum_{k=0}^{\infty} a_k(x-c)^k = \sum_{k=1}^{\infty} ka_k(x-c)^{k-1} \qquad \forall\, x \in (c-R,c+R)$$
and
$$\int_{\alpha}^{\beta}\sum_{k=0}^{\infty} a_k(x-c)^k dx = \sum_{k=0}^{\infty} a_k\int_{\alpha}^{\beta}(x-c)^k dx.$$

Example 7.39. Let $\{a_k\}_{k=0}^{\infty}$ be a bounded sequence. Then
$$\frac{d}{dx}\Big(\sum_{k=0}^{\infty}\frac{a_k}{k!}x^k\Big) = \sum_{k=1}^{\infty}\frac{a_k}{(k-1)!}x^{k-1} = \sum_{k=0}^{\infty}\frac{a_{k+1}}{k!}x^k.$$

Example 7.40. We show $\displaystyle\int_0^t e^x dx = e^t - 1$ as follows. By Theorem 7.23, $e^x = \sum\limits_{k=0}^{\infty} \dfrac{x^k}{k!}$ and the convergence is uniform on any bounded sets of $\mathbb{R}$; thus Corollary 7.38 implies that

$$\int_0^t e^x dx = \int_0^t \sum_{k=0}^{\infty} \frac{x^k}{k!} dx = \sum_{k=0}^{\infty} \int_0^t \frac{x^k}{k!} dx = \sum_{k=0}^{\infty} \frac{t^{k+1}}{(k+1)!} = \sum_{k=1}^{\infty} \frac{t^k}{k!} = e^t - 1 .$$

Example 7.41. $\dfrac{d}{dx}\Big(\sum\limits_{k=1}^{\infty} \dfrac{x^k}{k}\Big) = \sum\limits_{k=1}^{\infty} x^{k-1} = \sum\limits_{k=0}^{\infty} x^k$ for all $x \in (-1,1)$; thus

$$\frac{d}{dx}\Big(\sum_{k=1}^{\infty} \frac{x^k}{k}\Big) = \frac{1}{1-x} \qquad \forall\, x \in (-1,1) .$$

As a consequence,

$$\sum_{k=1}^{\infty} \frac{t^k}{k} = \int_0^t \frac{d}{dx}\Big(\sum_{k=1}^{\infty} \frac{x^k}{k}\Big) dx = -\log(1-t) \qquad \forall\, t \in (-1,1) . \tag{7.3.1}$$

Using the alternating series test, it is clear that the left-hand side of (7.3.1) converges at $t = -1$. What is the value of

$$-\sum_{k=1}^{\infty} \frac{(-1)^k}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots ?$$

Consider the partial sum $\dfrac{d}{dx}\Big(\sum\limits_{k=1}^{n} \dfrac{x^k}{k}\Big) = \sum\limits_{k=0}^{n-1} x^k = \dfrac{1-x^n}{1-x} = \dfrac{1}{1-x} - \dfrac{x^n}{1-x}$. Integrating both sides over $[-1,0]$,

$$\Big|\sum_{k=1}^{n} \frac{(-1)^k}{k} + \log 2\Big| \leqslant \int_{-1}^{0} \frac{|x|^n}{1-x} dx \leqslant \int_{-1}^{0} (-x)^n dx = \frac{1}{n+1} \to 0 \text{ as } n \to \infty;$$

thus

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots = \log 2 .$$

In other words,

$$\sum_{k=1}^{\infty} \frac{t^k}{k} = -\log(1-t) \qquad \forall\, t \in [-1,1) .$$

Example 7.42. It is clear that $\dfrac{1}{1+x^2} = \sum\limits_{k=0}^{\infty} (-x^2)^k = \sum\limits_{k=0}^{\infty} (-1)^k x^{2k}$ for all $x \in (-1,1)$. So if $x \in (-1,1)$,

$$\begin{aligned} \tan^{-1} x &= \int_0^x \frac{dt}{1+t^2} = \int_0^x \sum_{k=0}^{\infty} (-1)^k t^{2k} dt = \sum_{k=0}^{\infty} \int_0^x (-1)^k t^{2k} dt \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} t^{2k+1} \Big|_{t=0}^{t=x} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots . \end{aligned}$$

The right-hand side of the identity above converges at $x = 1$. What is the value of

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots ?$$

Mimic the previous example, we consider

$$\begin{aligned}
\tan^{-1} x &= \int_0^x \frac{dt}{1+t^2} = \int_0^x \frac{1-(-t^2)^{n+1}}{1+t^2} dt + \int_0^x \frac{(-t^2)^{n+1}}{1+t^2} dt \\
&= \int_0^x \sum_{k=0}^{n} (-1)^k t^{2k} dt + \int_0^x \frac{(-t^2)^{n+1}}{1+t^2} dt \\
&= \sum_{k=0}^{n} \int_0^x (-1)^k t^{2k} dt + \int_0^x \frac{(-t^2)^{n+1}}{1+t^2} dt = \sum_{k=0}^{n} \frac{(-1)^k}{2k+1} x^{2k+1} + \int_0^x \frac{(-t^2)^{n+1}}{1+t^2} dt;
\end{aligned}$$

thus plugging $x = 1$,

$$\Big| \tan^{-1} 1 - \sum_{k=0}^{n} \frac{(-1)^k}{2k+1} \Big| \leqslant \int_0^1 \frac{t^{2(n+1)}}{1+t^2} dt \leqslant \int_0^1 t^{2(n+1)} dt = \frac{1}{2n+3} \to 0 \text{ as } n \to \infty.$$

Therefore,

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \tan^{-1} 1 = \frac{\pi}{4}.$$

7.4 The Space of Continuous Functions

Definition 7.43. Let (M, d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $A \subseteq M$ be a subset. We define $\mathscr{C}(A; \mathcal{V})$ as the collection of all continuous functions on A with value in $\mathcal{V}$; that is,

$$\mathscr{C}(A; \mathcal{V}) = \big\{ f : A \to \mathcal{V} \,\big|\, f \text{ is continuous on } A \big\}.$$

Let $\mathscr{C}_b(A; \mathcal{V})$ be the subspace of $\mathscr{C}(A; \mathcal{V})$ which consists of all bounded continuous functions on A; that is,

$$\mathscr{C}_b(A; \mathcal{V}) = \big\{ f \in \mathscr{C}(A; \mathcal{V}) \,\big|\, f \text{ is bounded} \big\}.$$

Every $f \in \mathscr{C}_b(A; \mathcal{V})$ is associated with a non-negative real number $\|f\|_\infty$ given by

$$\|f\|_\infty = \sup \big\{ \|f(x)\| \,\big|\, x \in A \big\} = \sup_{x \in A} \|f(x)\|.$$

The number $\|f\|_\infty$ is called the ***sup-norm*** of f.

Proposition 7.44. *Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a normed vector space, $A \subseteq M$ be a subset.*

1. *$\mathscr{C}(A;\mathcal{V})$ and $\mathscr{C}_b(A;\mathcal{V})$ are vector spaces.*
2. *$\big(\mathscr{C}_b(A;\mathcal{V}),\|\cdot\|_\infty\big)$ is a normed vector space.*
3. *If $K \subseteq M$ is compact, then $\mathscr{C}(K;\mathcal{V}) = \mathscr{C}_b(K;\mathcal{V})$.*

Proof. 1 and 2 are trivial, and 3 is concluded by Theorem 4.25. ▫

Remark 7.45. In general $\|\cdot\|_\infty$ is not a "norm" on $\mathscr{C}(A;\mathcal{V})$. For example, the function $f(x) = \dfrac{1}{x}$ belongs to $\mathscr{C}((0,1);\mathbb{R})$ and $\|f\|_\infty = \infty$. Note that to be a norm $\|f\|_\infty$ has to take values in $\mathbb{R}$, and $\infty \notin \mathbb{R}$.

Proposition 7.46. *Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a normed vector space, $A \subseteq M$ be a subset, and $f_k \in \mathscr{C}_b(A;\mathcal{V})$ for all $k \in \mathbb{N}$. Then $\{f_k\}_{k=1}^\infty$ converges uniformly on A if and only if $\{f_k\}_{k=1}^\infty$ converges in $\big(\mathscr{C}_b(A;\mathcal{V}),\|\cdot\|_\infty\big)$.*

Proof. (⇐) Suppose that $\{f_k\}_{k=1}^\infty$ converges in $\big(\mathscr{C}_b(A;\mathcal{V}),\|\cdot\|_\infty\big)$. Then there exists $f \in \big(\mathscr{C}_b(A;\mathcal{V}),\|\cdot\|_\infty\big)$ such that $\lim\limits_{k\to\infty}\|f_k - f\|_\infty = 0$, and by the definition of the sup-norm,

$$\lim_{k\to\infty}\sup_{x\in A}\|f_k(x) - f(x)\| = 0\,.$$

Therefore, $\{f_k\}_{k=1}^\infty$ converges to f uniformly on A.

(⇒) Suppose that $\{f_k\}_{k=1}^\infty$ converges uniformly on A. Then there exists a function $f : A \to \mathcal{V}$ such that

$$\lim_{k\to\infty}\sup_{x\in A}\|f_k(x) - f(x)\| = 0\,.$$

By the definition of the sup-norm, it suffices to show that $f \in \mathscr{C}_b(A;\mathcal{V})$ in order to conclude the proposition. By Theorem 7.8, we obtain that $f \in \mathscr{C}(A;\mathcal{V})$. Moreover, the uniform convergence implies that there exists $N > 0$ such that

$$\|f_k(x) - f(x)\| < 1 \qquad \forall\, k \geqslant N \text{ and } x \in A\,.$$

In particular, the boundedness of f_N provides $M > 0$ such that $\|f_N(x)\| \leqslant M$ for all $x \in A$; thus

$$\|f(x)\| \leqslant \|f_N(x)\| + \|f(x) - f_N(x)\| \leqslant M + 1 \qquad \forall\, x \in A\,.$$

This implies that f is bounded; thus $f \in \mathscr{C}_b(A;\mathcal{V})$. ▫

Theorem 7.47. *Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a normed vector space, and $A \subseteq M$ be a subset. If $(\mathcal{V},\|\cdot\|)$ is complete, so is $\big(\mathscr{C}_b(A;\mathcal{V}),\|\cdot\|_\infty\big)$.*

Proof. Let $\{f_k\}_{k=1}^\infty$ be a Cauchy sequence in $\big(\mathscr{C}_b(A;\mathcal{V}),\|\cdot\|_\infty\big)$. Then

$$\forall\,\varepsilon>0, \exists\,N>0 \ni \|f_k-f_\ell\|_\infty<\varepsilon \quad \text{whenever} \quad k,\ell \geqslant N\,.$$

By the definition of the sup-norm, the statement above implies that

$$\forall\,\varepsilon>0, \exists\,N>0 \ni \|f_k(x)-f_\ell(x)\|<\varepsilon \quad \text{whenever} \quad k,\ell \geqslant N \text{ and } x\in A$$

which shows that $\{f_k\}_{k=1}^\infty$ converges uniformly on A because of the Cauchy criterion. Proposition 7.46 then implies that $\{f_k\}_{k=1}^\infty$ converges in $\big(\mathscr{C}_b(A;\mathcal{V}),\|\cdot\|_\infty\big)$. $\square$

Example 7.48. In this example we try to visualize a ball in $\mathscr{C}_b(A,\mathcal{V})$. Note that if $f\in\mathscr{C}_b(A,\mathcal{V})$ and $\varepsilon>0$. Then

$$B(f,\varepsilon)=\big\{g\in\mathscr{C}_b(A,\mathcal{V})\,\big|\,\|f-g\|_\infty<\varepsilon\big\}.$$

In particular, if $A=[a,b]$ and $\mathcal{V}=\mathbb{R}$, then $g\in B(f,\varepsilon)$ if and only if $|f(x)-g(x)|<\varepsilon$ for all $x\in[a,b]$ which means that the graph of g lies between the graph of $y=f(x)+\varepsilon$ and $y=f(x)-\varepsilon$.

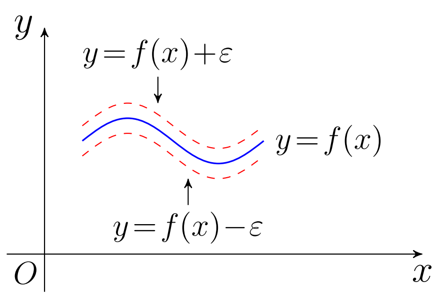


Figure 7.2: $g\in B(f,\varepsilon)$ if the graph of g lies in between the two red dash lines

Example 7.49. The set $B=\big\{f\in\mathscr{C}([0,1];\mathbb{R})\,\big|\,f(x)>0 \text{ for all } x\in[0,1]\big\}$ is open in $\big(\mathscr{C}([0,1];\mathbb{R}),\|\cdot\|_\infty\big)$.

Reason: Let $f\in B$ be given. Since $[0,1]$ is compact and f is continuous, by the extreme value theorem there exists $x_0\in[0,1]$ so that $\inf\limits_{x\in[0,1]} f(x)=f(x_0)>0$. Take $\varepsilon=\dfrac{f(x_0)}{2}$. Now if g is such that $\|g - f\|_\infty = \sup\limits_{x\in[0,1]} \big|g(x) - f(x)\big| < \varepsilon = \dfrac{f(x_0)}{2}$, we have for any $y \in [0,1]$,

$$\big|g(y) - f(y)\big| \leqslant \sup_{x\in[0,1]} \big|g(x) - f(x)\big| < \frac{f(x_0)}{2};$$

thus

$$g(y) \geqslant f(y) - \frac{f(x_0)}{2} \geqslant f(x_0) - \frac{f(x_0)}{2} = \frac{f(x_0)}{2} > 0.$$

Therefore, $g \in B$; thus $B(f, \varepsilon) \subseteq B$.

Example 7.50. Find the closure of B given in the previous example.

Proof. Claim: $\mathrm{cl}(B) = \big\{f \in \mathscr{C}([0,1],\mathbb{R}) \,\big|\, f(x) \geqslant 0\big\}$.
Proof of claim: We show that for every $f \in \big\{f \in \mathscr{C}([0,1],\mathbb{R}) \,\big|\, f(x) \geqslant 0\big\}$, there exists $f_k \in B$ such that $\|f_k - f\|_\infty \to 0$ as $k \to \infty$. Take $f_k(x) = f(x) + \dfrac{1}{k}$, then $f_k \in B$ ($\because f_k(x) > 0$), and

$$\|f_k - f\|_\infty = \sup_{x\in[0,1]} \big|f_k(x) - f(x)\big| \leqslant \sup_{x\in[0,1]} \frac{1}{k} = \frac{1}{k} \to 0 \text{ as } k \to \infty. \qquad \square$$

7.5 The Stone-Weierstrass Theorem

Theorem 7.51 (Weierstrass)**.** *Let $f : [0,1] \to \mathbb{R}$ be continuous. Then for every $\varepsilon > 0$, there exists a polynomial $p : [0,1] \to \mathbb{R}$ such that $\|f - p\|_\infty < \varepsilon$. In other words, the collection of all polynomials is dense in the space $\big(\mathscr{C}([0,1];\mathbb{R}), \|\cdot\|_\infty\big)$.*

Proof. For a fixed $n \in \mathbb{N}$, let $r_k(x) = C^n_k x^k (1-x)^{n-k}$. By looking at the partial derivatives with respect to x of the identity $(x+y)^n = \sum\limits_{k=0}^{n} C^n_k x^k y^{n-k}$, we find that

$$1.\ \sum_{k=0}^{n} r_k(x) = 1; \quad 2.\ \sum_{k=0}^{n} k r_k(x) = nx; \quad 3.\ \sum_{k=0}^{n} k(k-1) r_k(x) = n(n-1)x^2.$$

As a consequence,

$$\sum_{k=0}^{n} (k - nx)^2 r_k(x) = \sum_{k=0}^{n} \big[k(k-1) + (1 - 2nx)k + n^2x^2\big] r_k(x) = nx(1-x).$$

Let $\varepsilon > 0$ be given. Since $f : [0,1] \to \mathbb{R}$ is continuous on a compact $[0,1]$, f is uniformly continuous on $[0,1]$ (by Theorem 4.49); thus there exists $\delta > 0$ such that

$$\big|f(x) - f(y)\big| < \frac{\varepsilon}{2} \quad \text{whenever} \quad |x - y| < \delta,\ x, y \in [0,1].$$

Choose $n \in \mathbb{N}$ such that $\dfrac{\|f\|_\infty}{n\delta^2} < \varepsilon$, and define the ***Bernstein polynomial*** $p(x) = \sum\limits_{k=0}^{n} f\big(\frac{k}{n}\big) r_k(x)$. Then p is a polynomial. Moreover, for $x \in [0,1]$ we have

$$\begin{aligned}
\big|f(x) - p(x)\big| &= \Big|\sum_{k=0}^{n} \Big(f(x) - f\big(\frac{k}{n}\big)\Big) r_k(x)\Big| \leqslant \sum_{k=0}^{n} \Big|f(x) - f\big(\frac{k}{n}\big)\Big| r_k(x) \\
&\leqslant \Big(\sum_{|k-nx|<\delta n} + \sum_{|k-nx|\geqslant\delta n}\Big)\Big|f(x) - f\big(\frac{k}{n}\big)\Big| r_k(x) \\
&< \frac{\varepsilon}{2} + 2\|f\|_\infty \sum_{|k-nx|\geqslant\delta n} \frac{(k-nx)^2}{(k-nx)^2} r_k(x) \\
&\leqslant \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{n^2\delta^2} \sum_{k=0}^{n} (k-nx)^2 r_k(x) \leqslant \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{n\delta^2} x(1-x)\,.
\end{aligned}$$

Since $\sup\limits_{x\in[0,1]} x(1-x) = \dfrac{1}{4}$, we find that

$$\|f - p\|_\infty = \sup_{x\in[0,1]} \big|f(x) - p(x)\big| \leqslant \frac{\varepsilon}{2} + \frac{\|f\|_\infty}{2n\delta^2} < \varepsilon\,. \qquad \square$$

Remark 7.52. A polynomial of the form $p_n(x) = \sum\limits_{k=0}^{n} \beta_k r_k(x)$ is called a ***Bernstein polynomial of degree*** n, and the coefficients β_k are called Bernstein coefficients.

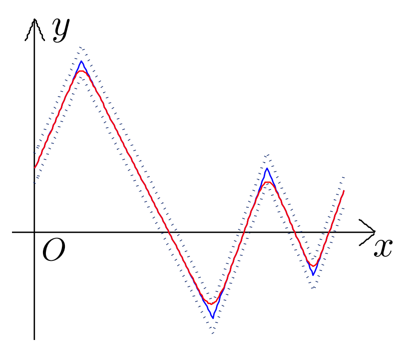


Figure 7.3: Using a Bernstein polynomial of degree 350 (the red curve) to approximate a "saw-tooth" function (the blue curve)

Corollary 7.53. *The collection of polynomials on* $[a,b]$ *is dense in* $\big(\mathscr{C}([a,b];\mathbb{R}), \|\cdot\|_\infty\big)$.

Proof. Let $g \in \mathscr{C}([a,b];\mathbb{R})$. Define $f(x) = g(x(b-a)+a)$. Then $f \in \mathscr{C}([0,1];\mathbb{R})$; thus there exists a sequence $p_n \in \mathscr{C}([0,1];\mathbb{R})$ such that

$$\lim_{n\to\infty} \sup_{y\in[0,1]} \big|f(y) - p_n(y)\big| = 0\,.$$

Therefore, with the change of variable $y = \dfrac{x-a}{b-a}$ (or $x = y(b-a)+a$),

$$\lim_{n\to\infty} \sup_{x\in[a,b]} \Big| g(x) - p_n\Big(\frac{x-a}{b-a}\Big)\Big| = \lim_{n\to\infty} \sup_{y\in[0,1]} \big|f(y) - p_n(y)\big| = 0;$$

thus by the fact $p_n\Big(\dfrac{x-a}{b-a}\Big)$ is a polynomial in x for all $n \in \mathbb{N}$ we conclude that there exists a sequence of polynomials converging to g uniformly on $[a,b]$. □

Definition 7.54. Let (M,d) be a metric space, and $E \subseteq M$ be a subset. A family $\mathcal{A}$ of real-valued functions defined on E is called an ***algebra*** if

1. $f + g \in \mathcal{A}$ for all $f, g \in \mathcal{A}$;
2. $f \cdot g \in \mathcal{A}$ for all $f, g \in \mathcal{A}$;
3. $\alpha f \in \mathcal{A}$ for all $f \in \mathcal{A}$ and $\alpha \in \mathbb{R}$.

In other words, $\mathcal{A}$ is an algebra if $\mathcal{A}$ is closed under addition, multiplication, and scalar multiplication.

Example 7.55. A function $g : [a,b] \to \mathbb{R}$ is called ***simple*** if we can divide up $[a,b]$ into sub-intervals on which g is constant except perhaps at the end-points. In other words, g is called simple if there is a partition $\mathcal{P} = \{x_0, x_1, \cdots, x_N\}$ of $[a,b]$ such that

$$g(x) = g\Big(\frac{x_{i-1}+x_i}{2}\Big) \qquad \text{if} \;\; x \in (x_{i-1}, x_i)\,.$$

Then the collection of all simple functions is an algebra.

Proposition 7.56. *Let (M,d) be a metric space, and $A \subseteq M$ be a subset. If $\mathcal{A} \subseteq \mathscr{C}_b(A;\mathbb{R})$ is an algebra, so is $\bar{\mathcal{A}}$.*

Proof. Let $f, g \in \bar{\mathcal{A}}$. Then there exists $\{f_k\}_{k=1}^\infty, \{g_k\}_{k=1}^\infty \subseteq \mathcal{A}$ such that $\{f_k\}_{k=1}^\infty$ converges uniformly to f on A, and $\{g_k\}_{k=1}^\infty$ converges uniformly to g on A. Since $\mathcal{A}$ is an algebra, $f_k + g_k, f_k \cdot g_k$ and αf_k belong to $\mathcal{A}$ for all $k \in \mathbb{N}$. By Theorem 2.48 and Proposition 7.46, the limit of $\{f_k + g_k\}_{k=1}^\infty$ and $\{\alpha f_k\}_{k=1}^\infty$ belong to $\bar{\mathcal{A}}$ which implies that $f+g$ and αf belong to $\bar{\mathcal{A}}$. Moreover,

$$\|f_k \cdot g_k - f \cdot g\|_\infty \leqslant \|f_k - f\|_\infty \|g_k\|_\infty + \|f\|_\infty \|g_k - g\|_\infty$$

which converges to 0 as $k \to \infty$; thus $f \cdot g$ is the limit of $\{f_k \cdot g_k\}_{k=1}^\infty$ so that $f \cdot g \in \bar{\mathcal{A}}$. Therefore, $\bar{\mathcal{A}}$ is an algebra. □

Corollary 7.57. *Let (M,d) be a metric space, $K \subseteq M$ be a compact set, and $\mathcal{A} \subseteq \mathscr{C}(K;\mathbb{R})$ be an algebra.*

1. *If $f \in \bar{\mathcal{A}}$, so is $|f|$.*

2. *If $f_1, \cdots, f_n \in \bar{\mathcal{A}}$, then $\max\{f_1,\cdots,f_n\} \in \bar{\mathcal{A}}$ and $\min\{f_1,\cdots,f_n\} \in \bar{\mathcal{A}}$, where*

$$\max\{f_1,\cdots,f_n\}(x) = \max\{f_1(x),\cdots,f_n(x)\},$$
$$\min\{f_1,\cdots,f_n\}(x) = \min\{f_1(x),\cdots,f_n(x)\}.$$

Proof. 1. Let $f \in \bar{\mathcal{A}}$. Then f is bounded so that $\mathrm{M} = \sup\limits_{x\in K}|f(x)| \in \mathbb{R}$. By Corollary 7.53, there exists a sequence of polynomial $\{p_n\}_{n=1}^\infty$ such that $\lim\limits_{n\to\infty} \sup\limits_{y\in[-M,M]} \big|p_n(y) - |y|\big| = 0$. Since $\mathcal{A}$ is an algebra, $\bar{\mathcal{A}}$ is also an algebra; thus $g_n \equiv p_n(f) \in \bar{\mathcal{A}}$. Moreover,

$$\sup_{x\in K}\big|g_n(x) - |f(x)|\big| = \sup_{x\in K}\big|p_n(f(x)) - |f(x)|\big| \leqslant \sup_{y\in[-M,M]}\big|p_n(y) - |y|\big|$$

which shows that $\{g_n\}_{n=1}^\infty$ converges uniformly to $|f|$ on K; thus $|f| \in \bar{\bar{\mathcal{A}}} = \bar{\mathcal{A}}$.

2. It suffices to show that $\max\{f,g\}$ and $\min\{f,g\}$ both belong to $\bar{\mathcal{A}}$ since

$$\max\{f_1,\cdots,f_n\} = \max\{\max\{f_1,\cdots,f_{n-1}\},f_n\},$$
$$\min\{f_1,\cdots,f_n\} = \min\{\min\{f_1,\cdots,f_{n-1}\},f_n\}.$$

Nevertheless, note that $\max\{f,g\} = \dfrac{f+g}{2} + \dfrac{|f-g|}{2}$ and $\min\{f,g\} = \dfrac{f+g}{2} - \dfrac{|f-g|}{2}$, we find that if $f,g \in \bar{\mathcal{A}}$ then $\max\{f,g\} \in \bar{\mathcal{A}}$ and $\min\{f,g\} \in \bar{\mathcal{A}}$. □

Definition 7.58. Let (M,d) be a metric space, and $E \subseteq M$ be a subset. A family $\mathscr{F}$ of real-valued functions defined on E is said to

1. ***separate points*** on E if for all $x,y \in E$ and $x \neq y$, there exists $f \in \mathscr{F}$ such that $f(x) \neq f(y)$.

2. ***vanish at no point*** of E if for each $x \in E$ there is $f \in \mathscr{F}$ such that $f(x) \neq 0$.

Example 7.59. Let $\mathscr{P}([a,b])$ denote the collection of polynomials defined on $[a,b]$ is an algebra. Moreover, $\mathscr{P}([a,b])$ separates points on $[a,b]$ since $p(x) = x$ does the separation, and $\mathscr{P}([a,b])$ vanishes at no point of $[a,b]$.

Example 7.60. Let $\mathscr{P}_{\text{even}}([a,b])$ denote the collection of all polynomials $p(x)$ of the form

$$p(x) = \sum_{k=0}^{n} a_k x^{2k} = a_n x^{2n} + a_{n-1} x^{2n-2} + \cdots + a_0 \,.$$

Then $\mathscr{P}_{\text{even}}([a,b])$ is an algebra. Moreover, $\mathscr{P}_{\text{even}}([a,b])$ vanishes at no point of $[a,b]$ since the constant functions are polynomials (since constant functions belongs to $\mathscr{P}([a,b])$. However, if $ab < 0$, $\mathscr{P}_{\text{even}}([a,b])$ does not separate points on $[a,b]$. On the other hand, if $ab \geqslant 0$, then $\mathscr{P}_{\text{even}}([a,b])$ separates points on $[a,b]$ since $p(x) = x^2$ does the job.

Lemma 7.61. *Let (M,d) be a metric space, and $E \subseteq M$ be a subset. Suppose that $\mathcal{A} \subseteq \mathscr{C}_b(E;\mathbb{R})$ is an algebra, $\mathcal{A}$ separates points on E, and $\mathcal{A}$ vanishes at no point of E. Then for all $x_1, x_2 \in E$, $x_1 \neq x_2$, and $c_1, c_2 \in \mathbb{R}$ (c_1, c_2 could be the same), there exists $f \in \mathcal{A}$ such that $f(x_1) = c_1$ and $f(x_2) = c_2$.*

Proof. Since $\mathcal{A}$ separates points on E, there exists $g \in \mathcal{A}$ such that $g(x_1) \neq g(x_2)$, and since $\mathcal{A}$ vanishes at no point of E, there exists $h, k \in \mathcal{A}$ such that $h(x_1) \neq 0$ and $k(x_2) \neq 0$. Then

$$f(x) = c_1 \frac{\big[g(x) - g(x_2)\big]h(x)}{\big[g(x_1) - g(x_2)\big]h(x_1)} + c_2 \frac{\big[g(x) - g(x_1)\big]k(x)}{\big[g(x_2) - g(x_1)\big]k(x_2)}$$

has the desired property. □

Theorem 7.62 (Stone)**.** *Let (M,d) be a metric space, $K \subseteq M$ be a compact set, and $\mathcal{A} \subseteq \mathscr{C}(K;\mathbb{R})$ satisfying*

1. *$\mathcal{A}$ is an algebra.* 2. *$\mathcal{A}$ separates points on K.* 3. *$\mathcal{A}$ vanishes at no point of K.*

Then $\mathcal{A}$ is dense in $\mathscr{C}(K;\mathbb{R})$; that is, for every $f \in \mathscr{C}(K;\mathbb{R})$ and $\varepsilon > 0$, there exists $g \in \mathcal{A}$ such that $\|f - g\|_\infty < \varepsilon$.

Proof. We first show that for any given $f \in \mathscr{C}(K;\mathbb{R})$, $a \in K$ and $\varepsilon > 0$, there exists a function $g_a \in \bar{\mathcal{A}}$ such that

$$g_a(a) = f(a) \qquad \text{and} \qquad g_a(x) > f(x) - \varepsilon \quad \forall\, x \in K \,. \tag{7.5.1}$$

Let $f \in \mathscr{C}(K;\mathbb{R})$, $a \in K$ and $\varepsilon > 0$ be given. Since $\mathcal{A}$ is an algebra, so is $\bar{\mathcal{A}}$; thus Lemma 7.61 implies that there exists $h_b \in \bar{\mathcal{A}}$ such that $h_b(a) = f(a)$ and $h_b(b) = f(b)$. Note that every function in $\bar{\mathcal{A}}$ is continuous (by Theorem 7.8); thus the continuity of h_b provides $\delta = \delta_b > 0$ such that

$$h_b(x) > f(x) - \varepsilon \qquad \forall\, x \in \big[B\big(b,\delta_b\big) \cup B\big(a,\delta_b\big)\big] \cap K \,.$$

Let $U_b = B\big(b, \delta_b\big) \cup B\big(a, \delta_b\big)$. Then U_b is open. Since $K \subseteq \bigcup\limits_{\substack{b \in K \\ b \neq a}} U_b$ and K is compact, there exists a finite set $\{b_1, \cdots, b_m\} \subseteq K \backslash \{a\}$ such that $K \subseteq \bigcup\limits_{j=1}^{n} U_{b_j}$. Define $g_a = \max\big\{h_{b_1}, \cdots h_{b_m}\big\}$. Then $g_a(a) = f(a)$, and Corollary 7.57 implies that $g_a \in \bar{\mathcal{A}}$. Moreover, if $x \in K$, $x \in U_{b_j}$ for some j; thus

$$g_a(x) \geqslant h_{b_j}(x) > f(x) - \varepsilon$$

which implies that g satisfies (7.5.1).

Let $f \in \mathscr{C}(K; \mathbb{R})$ and $\varepsilon > 0$ be given. For any $a \in K$, let $g_a \in \bar{\mathcal{A}}$ be a function satisfying

$$g_a(a) = f(a) \quad \text{and} \quad g_a(x) > f(x) - \frac{\varepsilon}{2} \qquad \forall\, x \in K\,. \tag{7.5.2}$$

By the continuity of g_a, there exists $\delta = \delta_a > 0$ such that

$$g_a(x) < f(x) + \frac{\varepsilon}{2} \qquad \forall\, x \in B(a, \delta_a) \cap K\,. \tag{7.5.3}$$

By the compactness of K, there exists $\{a_1, \cdots, a_n\} \subseteq K$ such that

$$K \subseteq \bigcup_{j=1}^{m} B\big(a_j, \delta_{a_j}\big)\,.$$

Define $h = \min\big\{g_{a_1}, \cdots, g_{a_n}\big\}$. Corollary 7.57 implies that $h \in \bar{\mathcal{A}}$, and (7.5.2) shows that

$$h(x) > f(x) - \frac{\varepsilon}{2} \qquad \forall\, x \in K.$$

Moreover, if $x \in K$, there exists j such that $x \in B\big(a_j, \delta_{a_j}\big)$ and (7.5.3) further shows that

$$h(x) \leqslant g_{a_j}(x) < f(x) + \frac{\varepsilon}{2};$$

thus

$$h(x) < f(x) + \frac{\varepsilon}{2} \qquad \forall\, x \in K.$$

Therefore, we establish the existence of $h \in \bar{\mathcal{A}}$ such that

$$\big|h(x) - f(x)\big| < \frac{\varepsilon}{2} \qquad \forall\, x \in K\,.$$

On the other hand, since $h \in \bar{\mathcal{A}}$, there exists $p \in \mathcal{A}$ such that

$$\big|p(x) - h(x)\big| < \frac{\varepsilon}{2} \qquad \forall\, x \in K\,;$$

thus

$$\big|p(x) - f(x)\big| \leqslant \big|p(x) - h(x)\big| + \big|h(x) - f(x)\big| < \varepsilon \qquad \forall\, x \in K$$

which concludes the theorem. □

Example 7.63. Let $K = [-1,1] \times [-1,1] \subseteq \mathbb{R}^2$. Consider the set $\mathscr{P}(K)$ of all polynomials $p(x, y)$ in two variables $(x, y) \in K$. Then $\mathscr{P}(K)$ is dense in $\mathscr{C}(K;\mathbb{R})$.

Reason: Since K is compact, and $\mathscr{P}(K)$ is definitely an algebra and the constant function $p(x, y) = 1 \in \mathscr{P}(K)$ vanishes at no point of K, it suffices to show that $\mathscr{P}(K)$ separates points. Let (a_1, b_1) and (a_2, b_2) be two different points in K. Then the polynomial

$$p(x, y) = (x - a_1)^2 + (y - b_1)^2$$

has the property that $p(a_1, b_1) \neq p(a_2, b_2)$. Therefore, $\mathscr{P}(K)$ separates points in K,

Example 7.64. Consider $\mathscr{P}_{\text{even}}([0,1]) = \Big\{p(x) = \sum\limits_{k=0}^{n} a_k x^{2k} \,\Big|\, a_k \in \mathbb{R}\Big\}$ (see Example 7.60). Then $\mathcal{A} = \mathscr{P}_{\text{even}}([0,1])$ satisfies all the conditions in the Stone theorem, so $\mathscr{P}_{\text{even}}([0,1])$ is dense in $\mathscr{C}([0,1];\mathbb{R})$.

On the other hand, if $K = [-1,1]$, then $\mathscr{P}_{\text{even}}([-1,1])$ does not separate points on K since if $p \in \mathscr{P}_{\text{even}}([-1,1])$, $p(x) = p(-x)$; thus the Stone theorem cannot be applied to conclude the denseness of $\mathscr{P}_{\text{even}}([-1,1])$ in $\mathscr{C}([-,1];\mathbb{R})$. In fact, $\mathscr{P}_{\text{even}}([-1,1])$ is not dense in $\mathscr{C}([-1,1];\mathbb{R})$ since polynomials in $\mathscr{P}_{\text{even}}([-1,1])$ are all even functions and $f(x) = x$ cannot be approximated by even functions.

Corollary 7.65. *Let $\mathscr{C}(\mathbb{T})$ be the collection of all 2π-periodic continuous real-valued functions, and $\mathscr{P}_n(\mathbb{T})$ be the collection of all real-valued trigonometric polynomials of degree n; that is,*

$$\mathscr{P}_n(\mathbb{T}) = \Big\{\frac{c_0}{2} + \sum_{k=1}^{n} c_k \cos kx + s_k \sin kx \,\Big|\, \{c_k\}_{k=0}^{n}, \{s_k\}_{k=1}^{n} \subseteq \mathbb{R}\Big\}.$$

Then $\mathscr{P}(\mathbb{T}) \equiv \bigcup\limits_{n=0}^{\infty} \mathscr{P}_n(\mathbb{T})$ is dense in $\mathscr{C}(\mathbb{T})$. In other words, if $f \in \mathscr{C}(\mathbb{T})$ and $\varepsilon > 0$ is given, there exists $p \in \mathscr{P}(\mathbb{T})$ such that

$$\big|f(x) - p(x)\big| < \varepsilon \qquad \forall\, x \in \mathbb{R}.$$

Proof. We note that $\mathscr{C}(\mathbb{T})$ can be viewed as the collection of all continuous functions defined on the unit circle $\mathbb{S}^1$ in the sense that every $f \in \mathscr{C}(\mathbb{T})$ corresponds to a unique $F \in \mathscr{C}(\mathbb{S}^1;\mathbb{R})$ such that $f(x) = F(\cos x, \sin x)$, and vice versa. Since $\mathbb{S}^1 \subseteq [-1,1] \times [-1,1]$ is compact, Example 7.63 provides that $\mathscr{P}(\mathbb{S}^1)$, the collection of all polynomials defined on $\mathbb{S}^1$, is an algebra that separates points of $\mathbb{S}^1$ and vanishes at no point on $\mathbb{S}^1$. The Stone-Weierstrass Theorem then implies that there exists $P \in \mathscr{P}(\mathbb{S}^1)$ such that

$$\big|F(x, y) - P(x, y)\big| < \varepsilon \qquad \forall\, (x, y) \in \mathbb{S}^1 \text{ (that is, } x^2 + y^2 = 1).$$

Let $p(x) = P(\cos x, \sin x)$. Note that

$$\begin{aligned}\cos^n x &= \Big(\frac{e^{ix}+e^{-ix}}{2}\Big)^n = \sum_{k=0}^n \frac{1}{2^n} C^n_k e^{ikx} e^{-i(n-k)x} = \sum_{k=0}^n \frac{1}{2^n} C^n_k e^{i(2k-n)x} \\ &= \sum_{k=0}^n \frac{1}{2^n} C^n_k \big(\cos(2k-n)x + i\sin(2k-n)x\big) = \sum_{k=0}^n \frac{1}{2^n} C^n_k \cos(2k-n)x \in \mathscr{P}_n(\mathbb{T}),\end{aligned}$$

and similarly, $\sin^m x \in \mathscr{P}_m(\mathbb{T})$. Therefore, if $P(x,y) = \sum\limits_{k,\ell=0}^n a_{k,\ell} x^k y^\ell$, then $P(\cos x, \sin x) \in \mathscr{P}_{2n}(\mathbb{T})$ because of the product-to-sum formulas

$$\begin{aligned}\cos\theta\cos\varphi &= \frac{1}{2}\Big[\cos(\theta-\varphi) + \cos(\theta+\varphi)\Big], \\ \sin\theta\cos\varphi &= \frac{1}{2}\Big[\sin(\theta+\varphi) + \sin(\theta-\varphi)\Big], \\ \sin\theta\sin\varphi &= \frac{1}{2}\Big[\cos(\theta-\varphi) - \cos(\theta+\varphi)\Big].\end{aligned}$$

As a consequence, we conclude that

$$\big|f(x) - p(x)\big| = \big|F(\cos x, \sin x) - P(\cos x, \sin x)\big| < \varepsilon \qquad \forall\, x \in \mathbb{R}.$$

□

7.6 Universal Approximation Theorem

The Stone–Weierstrass Theorem shows that continuous functions can be uniformly approximated by algebraically simple functions, such as polynomials. It is natural to ask whether other families of simple functions can also approximate continuous functions effectively. In this section we study approximation of continuous functions by linear combinations of ridge functions generated by a fixed nonlinearity. This viewpoint leads naturally to one-hidden-layer neural networks and to the Universal Approximation Theorem.

Ridge functions are functions that are constant on affine hyperplanes. More precisely, a ridge function on $\mathbb{R}^n$ is a function of the form

$$F(\boldsymbol{x}) = f(\boldsymbol{a}\cdot\boldsymbol{x} + b),$$

where $f : \mathbb{R} \to \mathbb{R}$, $\boldsymbol{a} \in \mathbb{R}^n \backslash \{0\}$, $b \in \mathbb{R}$, and

$$\boldsymbol{a}\cdot\boldsymbol{x} = \sum_{i=1}^n a_i x_i.$$

Since $f(\boldsymbol{a}\cdot\boldsymbol{x}+b)=\widetilde{f}(\boldsymbol{a}\cdot\boldsymbol{x})$ with $\widetilde{f}(t):=f(t+b)$, this is equivalent to the simpler form $f(\boldsymbol{a}\cdot\boldsymbol{x})$.

Recall that $\mathscr{C}(A;\mathbb{R}^m)$ denotes the space of continuous functions from a set A to $\mathbb{R}^m$.

Let $\Omega\subseteq\mathbb{R}^n$. We consider the linear space generated by ridge functions with directions in Ω:

$$\mathcal{M}(\Omega):=\operatorname{span}\left\{f(\boldsymbol{a}\cdot\boldsymbol{x})\,\middle|\,\boldsymbol{a}\in\Omega,\ f\in\mathscr{C}(\mathbb{R};\mathbb{R})\right\}.$$

For $k\in\mathbb{N}$, let H_k^n denote the space of homogeneous polynomials of degree k in n variables:

$$H_k^n=\left\{\sum_{|\boldsymbol{m}|=k}c_{\boldsymbol{m}}\boldsymbol{x}^{\boldsymbol{m}}\,\middle|\,\boldsymbol{m}\in\mathbb{Z}_+^n,\ c_{\boldsymbol{m}}\in\mathbb{R}\right\}.$$

Here we use multi-index notation: $\mathbb{Z}_+:=\mathbb{N}\cup\{0\}$ denotes the set of non-negative integers, $\boldsymbol{m}=(m_1,\ldots,m_n)\in\mathbb{Z}_+^n$, $|\boldsymbol{m}|=m_1+\cdots+m_n$, and

$$\boldsymbol{x}^{\boldsymbol{m}}=x_1^{m_1}\cdots x_n^{m_n}.$$

It is well known that H_k^n is finite dimensional, with

$$\dim H_k^n=\binom{n-1+k}{k}.$$

We also define

$$H^n:=\bigcup_{k=0}^{\infty}H_k^n,$$

the collection of all homogeneous polynomials in n variables.

Definition 7.66. Let X be a collection of real-valued functions on $\mathbb{R}^n$. We say that X is *dense in* $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$ in the sense of uniform convergence on compact subsets, or simply X is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$, if $X\subseteq\mathscr{C}(\mathbb{R}^n;\mathbb{R})$ and for every $\varepsilon>0$, $f\in\mathscr{C}(\mathbb{R}^n;\mathbb{R})$ and compact $K\subseteq\mathbb{R}^n$, there exists $g\in X$ such that

$$\sup_{x\in K}|f(x)-g(x)|<\varepsilon.$$

Proposition 7.67. *The space $\mathcal{M}(\Omega)$ is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$ (in the sense of uniform convergence on compact subsets) if and only if the only polynomial in H^n that vanishes identically on Ω is the zero polynomial.*

Proof. The direction "⇒" will not be used in this chapter. Its proof involves Fourier transform techniques and will be given as a guided exercise in Chapter 9 (see Problem 9.46).

We prove the direction "⇐". For $\boldsymbol{m} = (m_1, \ldots, m_n) \in \mathbb{Z}_+^n$ with $|\boldsymbol{m}| = k$, write

$$D^{\boldsymbol{m}} = \frac{\partial^k}{\partial x_1^{m_1} \cdots \partial x_n^{m_n}}.$$

Step 1: Representation of H_k^n.

We claim that

$$H_k^n = \operatorname{span}\left\{(\boldsymbol{a} \cdot \boldsymbol{x})^k \,\middle|\, \boldsymbol{a} \in \Omega\right\}.$$

The inclusion "⊇" is clear. Suppose, by contradiction, that the span is strictly smaller than H_k^n. Since H_k^n is finite-dimensional, there exists a nonzero linear functional $\ell : H_k^n \to \mathbb{R}$ such that

$$\ell(p) = 0 \quad \text{for all } p \in \operatorname{span}\left\{(\boldsymbol{a} \cdot \boldsymbol{x})^k \,\middle|\, \boldsymbol{a} \in \Omega\right\}.$$

Define

$$c_{\boldsymbol{m}} = \frac{1}{\boldsymbol{m}!}\,\ell(\boldsymbol{x}^{\boldsymbol{m}}), \qquad q(\boldsymbol{x}) = \sum_{|\boldsymbol{m}|=k} c_{\boldsymbol{m}} \boldsymbol{x}^{\boldsymbol{m}} \in H_k^n.$$

Then $q \neq 0$ by the nontriviality of ℓ. Using the identity

$$D^{\boldsymbol{m}_1} \boldsymbol{x}^{\boldsymbol{m}_2} = \delta_{\boldsymbol{m}_1, \boldsymbol{m}_2}\, \boldsymbol{m}_1!,$$

we obtain

$$\ell(p) = q(D)p \qquad \forall\, p \in H_k^n. \tag{7.6.1}$$

Now observe that for any $\boldsymbol{d} \in \mathbb{R}^n$,

$$D^{\boldsymbol{m}}(\boldsymbol{d} \cdot \boldsymbol{x})^k = k!\, \boldsymbol{d}^{\boldsymbol{m}}.$$

Therefore,

$$0 = \ell\big((\boldsymbol{a} \cdot \boldsymbol{x})^k\big) = q(D)(\boldsymbol{a} \cdot \boldsymbol{x})^k = k!\, q(\boldsymbol{a}) \qquad \forall\, \boldsymbol{a} \in \Omega.$$

Hence q vanishes identically on Ω. By assumption, this implies $q = 0$, a contradiction. Thus

$$H_k^n = \operatorname{span}\left\{(\boldsymbol{a} \cdot \boldsymbol{x})^k \,\middle|\middle|\, \boldsymbol{a} \in \Omega\right\}.$$

Step 2: Density.

Since $(\boldsymbol{a} \cdot \boldsymbol{x})^k \in \mathcal{M}(\Omega)$, we conclude that $H_k^n \subseteq \mathcal{M}(\Omega)$ for every k. Hence $\mathcal{M}(\Omega)$ contains all polynomials. By the Stone–Weierstrass Theorem, polynomials are dense in $\mathscr{C}(\mathbb{R}^n; \mathbb{R})$ under uniform convergence on compact sets. Therefore $\mathcal{M}(\Omega)$ is dense. □

Corollary 7.68. *The space*

$$V = \operatorname{span}\left\{ f(\boldsymbol{a}\cdot\boldsymbol{x}) \,\middle|\, \boldsymbol{a}\in\mathbb{R}^n,\ f\in\mathscr{C}(\mathbb{R};\mathbb{R})\right\}$$

is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$.

Proof. Apply Proposition 7.67 with $\Omega = \mathbb{R}^n$. In this case no nonzero polynomial can vanish on Ω. □

For $\sigma : \mathbb{R} \to \mathbb{R}$ and $n \in \mathbb{N}$, define

$$\Sigma_n(\sigma) = \operatorname{span}\left\{ \sigma(\boldsymbol{w}\cdot\boldsymbol{x}+\theta) \,\middle|\, \boldsymbol{w}\in\mathbb{R}^n,\ \theta\in\mathbb{R}\right\}. \tag{7.6.2}$$

Corollary 7.69. *Let* $\sigma \in \mathscr{C}(\mathbb{R};\mathbb{R})$. *If* $\Sigma_1(\sigma)$ *is dense in* $\mathscr{C}(\mathbb{R};\mathbb{R})$, *then* $\Sigma_n(\sigma)$ *is dense in* $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$.

Proof. Let $\varepsilon > 0$, $f \in \mathscr{C}(\mathbb{R}^n;\mathbb{R})$ and compact $K \subseteq \mathbb{R}^n$ be given.

By Corollary 7.68, there exist $g_i \in \mathscr{C}(\mathbb{R};\mathbb{R})$ and $\boldsymbol{a}_i \in \mathbb{R}^n$ such that

$$\left| f(\boldsymbol{x}) - \sum_{i=1}^{k} g_i(\boldsymbol{a}_i\cdot\boldsymbol{x}) \right| < \frac{\varepsilon}{2} \qquad \forall\, \boldsymbol{x}\in K.$$

For each i, $\boldsymbol{a}_i\cdot\boldsymbol{x}$ ranges in a compact interval $[\alpha_i,\beta_i]$ when $\boldsymbol{x}\in K$. Since $\Sigma_1(\sigma)$ is dense, there exist coefficients $c_{ij}, w_{ij}, \theta_{ij}$, $1 \leqslant j \leqslant m_i$, such that

$$\left| g_i(y) - \sum_{j=1}^{m_i} c_{ij}\sigma(w_{ij}y+\theta_{ij}) \right| < \frac{\varepsilon}{2k} \qquad \forall\, y\in[\alpha_i,\beta_i].$$

Setting $\boldsymbol{w}_{ij} = w_{ij}\boldsymbol{a}_i$, we obtain

$$\left| f(\boldsymbol{x}) - \sum_{i=1}^{k}\sum_{j=1}^{m_i} c_{ij}\sigma(\boldsymbol{w}_{ij}\cdot\boldsymbol{x}+\theta_{ij}) \right| < \varepsilon \qquad \forall\, \boldsymbol{x}\in K.$$

The double sum belongs to $\Sigma_n(\sigma)$, so $\Sigma_n(\sigma)$ is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$. □

The results obtained above show that linear combinations of ridge functions are sufficiently rich to approximate arbitrary continuous functions on $\mathbb{R}^n$, provided that the set of directions is large enough. In particular, Corollary 7.68 tells us that finite linear combinations of functions of the form $f(\boldsymbol{a}\cdot\boldsymbol{x})$ are dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$ under uniform convergence on compact sets.

This observation is closely related to neural networks. Indeed, a one-hidden-layer neural network with activation function σ produces functions of the form

$$\boldsymbol{x} \mapsto \sum_{j=1}^{k} c_j\, \sigma(\boldsymbol{a}_j \cdot \boldsymbol{x} + b_j),$$

which are precisely finite linear combinations of ridge functions. Thus the density of such networks is intimately connected with the approximation properties of ridge functions.

The general density theorem for ridge functions (Proposition 7.67) reduces the problem to a one-dimensional approximation question. More precisely, if we know that

$$\Sigma_1(\sigma) = \operatorname{span}\left\{\sigma(wx+\theta) \,\middle|\, w,\theta \in \mathbb{R}\right\}$$

is dense in $\mathscr{C}(\mathbb{R};\mathbb{R})$, then Corollary 7.69 implies that $\Sigma_n(\sigma)$ is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$ for every $n \in \mathbb{N}$.

We are therefore led to the following central question:

For which activation functions σ is the space $\Sigma_1(\sigma)$ dense in $\mathscr{C}(\mathbb{R};\mathbb{R})$?

The Universal Approximation Theorem provides a complete answer.

In the following, for a function $\sigma : \mathbb{R} \to \mathbb{R}$ and $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we define

$$(\sigma \circ \boldsymbol{x})_i := \sigma(x_i), \qquad i = 1, \ldots, n.$$

Thus $\sigma \circ \boldsymbol{x}$ denotes the vector obtained by applying σ to each component of $\boldsymbol{x}$. For notational simplicity, we shall also write

$$\sigma(\boldsymbol{x})$$

instead of $\sigma \circ \boldsymbol{x}$ whenever $\boldsymbol{x} \in \mathbb{R}^n$, with the understanding that σ acts componentwise.

Definition 7.70 (One-hidden-layer neural networks)**.** Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a function. For $n, m \in \mathbb{N}$, we define the class of ***one-hidden-layer neural networks*** with input dimension n and output dimension m by

$$\mathcal{N}_\sigma(n,m) := \left\{\boldsymbol{x} \mapsto C\sigma(A\boldsymbol{x} + \boldsymbol{b}) \,\middle|\, k \in \mathbb{N},\ A \in \mathbb{R}^{k\times n},\ \boldsymbol{b} \in \mathbb{R}^k,\ C \in \mathbb{R}^{m\times k}\right\}.$$

Remark 7.71.

1. In machine learning terminology, σ is called the ***activation function***. In the representation $C\sigma(A\boldsymbol{x}+\boldsymbol{b})$, the matrices A and C are called the ***weight matrices***, and $\boldsymbol{b}$ the ***bias vector***. The activation function provides the only source of nonlinearity in the network.

2. If $\sigma \in \mathscr{C}(\mathbb{R};\mathbb{R})$, then
$$\mathcal{N}_\sigma(n,m) \subseteq \mathscr{C}(\mathbb{R}^n;\mathbb{R}^m).$$

3. In the scalar-output case $m = 1$, every element of $\mathcal{N}_\sigma(n,1)$ has the form
$$g(\boldsymbol{x}) = \sum_{j=1}^{k} c_j\, \sigma(\boldsymbol{a}_j \cdot \boldsymbol{x} + b_j),$$
where $\boldsymbol{a}_j \in \mathbb{R}^n$ and $b_j, c_j \in \mathbb{R}$. Thus
$$\mathcal{N}_\sigma(n,1) = \Sigma_n(\sigma),$$
where $\Sigma_n(\sigma)$ was introduced in (7.6.2).

Theorem 7.72 (Universal Approximation Theorem)**.** *Let $\sigma \in \mathscr{C}(\mathbb{R};\mathbb{R})$. Then σ is not a polynomial if and only if $\mathcal{N}_\sigma(n,m)$ is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R}^m)$ (in the sense of uniform convergence on compact subsets).*

Proof. It suffices to prove the case $m = 1$, since uniform convergence of vector-valued functions is equivalent to uniform convergence of each component.

"$\Leftarrow$" Assume that σ is a polynomial of degree d, and let $K \subseteq \mathbb{R}^n$ be compact. Each function $\sigma(\boldsymbol{a}\cdot\boldsymbol{x}+b)$ is then a polynomial in $\boldsymbol{x}$ of degree at most d. Hence every element of $\mathcal{N}_\sigma(n,1)$ is a polynomial of degree at most d.

Define
$$\mathscr{P}_d(K) := \left\{\, p|_K \,\middle|\, p \in \mathscr{P}_d(\mathbb{R}^n) \,\right\} \subseteq \mathscr{C}(K;\mathbb{R}).$$
Then $\mathcal{N}_\sigma(n,1) \subseteq \mathscr{P}_d(K)$. Since $\mathscr{P}_d(K)$ is finite dimensional, it is closed in $\mathscr{C}(K;\mathbb{R})$. Thus
$$\overline{\mathcal{N}_\sigma(n,1)}^{\|\cdot\|_\infty} \subseteq \mathscr{P}_d(K).$$

If $\mathcal{N}_\sigma(n,1)$ were dense in $\mathscr{C}(K;\mathbb{R})$, we would have $\mathscr{P}_d(K) = \mathscr{C}(K;\mathbb{R})$, which is impossible since the former is finite dimensional while the latter is infinite dimensional. Hence density fails.

"$\Rightarrow$" Assume that σ is not a polynomial. Recall that $\mathcal{N}_\sigma(n,1) = \Sigma_n(\sigma)$, by Corollary 7.69 it suffices to prove density in dimension $n = 1$.

To remain within the scope of this chapter, we give a complete proof only for the ReLU activation

$$\sigma(t) = \max\{0, t\}.$$

The general case will be discussed in Remark 7.73.

Let $\varepsilon > 0$, $f \in \mathscr{C}(\mathbb{R};\mathbb{R})$, and compact $K \subseteq \mathbb{R}$ be given. Since K is compact, we can choose $L > 0$ such that $K \subseteq [-L, L]$. By the fact that f is uniformly continuous on $[-L, L]$, there exists $\delta > 0$ such that

$$\big|f(x_1) - f(x_2)\big| < \varepsilon \qquad \text{whenever} \quad |x_1 - x_2| < \delta \text{ and } x_1, x_2 \in [-L, L].$$

Let $\{-L = t_0 < t_1 < \cdots < t_M = L\}$ be a uniform partition of $[-L, L]$ with mesh size $\|P\| = \dfrac{2L}{M} < \delta$. Let p be the continuous piecewise linear function interpolating f at the partition points:

$$p(t_k) = f(t_k).$$

Let

$$s_k = \frac{f(t_k) - f(t_{k-1})}{t_k - t_{k-1}}$$

be the slope on (t_{k-1}, t_k). A direct computation shows

$$p(x) = p(t_0) + s_1(x - t_0) + \sum_{k=2}^{M} (s_k - s_{k-1})\, \sigma(x - t_{k-1}).$$

Indeed, defining

$$q(x) = p(t_0) + s_1(x - t_0) + \sum_{k=2}^{M} (s_k - s_{k-1})\, \sigma(x - t_{k-1}),$$

one verifies that q and p have the same slope on each interval (t_{k-1}, t_k) and agree at t_0, hence $q \equiv p$.

Since any affine function satisfies

$$ax + b = \sigma(ax + b) - \sigma(-ax - b),$$

we conclude that $p \in \mathcal{N}_\sigma(1, 1)$.

Finally, for $x \in [t_{k-1}, t_k]$,

$$p(x) = \frac{x - t_{k-1}}{t_k - t_{k-1}} f(t_k) + \frac{t_k - x}{t_k - t_{k-1}} f(t_{k-1}),$$

and a convex combination argument yields

$$\big|f(x) - p(x)\big| < \varepsilon \qquad \forall\, x \in [t_{k-1}, t_k], 1 \leqslant k \leqslant M.$$

Thus

$$\sup_{x \in K} |f(x) - p(x)| < \varepsilon.$$

Since $p \in \mathcal{N}_\sigma(1, 1)$, density follows. □

Remark 7.73. A complete proof of the Universal Approximation Theorem for general non-polynomial activations typically relies on tools from functional analysis and measure theory, such as:

- the Hahn–Banach separation theorem,
- the Riesz representation theorem for $\mathscr{C}(K; \mathbb{R})^*$,
- measure-theoretic arguments and convergence theorems (e.g. dominated convergence),
- in some approaches, Fourier or harmonic analysis techniques.

These topics go beyond the scope of this chapter on uniform convergence. For complete treatments, see [1, 2, 3, 4]. The sigmoidal case illustrates why functional analytic tools such as Hahn−Banach and the Riesz representation theorem naturally arise in this problem. See guided exercise problem 7.25.

The Universal Approximation Theorem shows that a single hidden layer suffices to approximate arbitrary continuous functions on compact sets. In this sense, no additional depth is required for universality.

However, the theorem is purely existential. It does not address how large the number of neurons must be in order to achieve a prescribed accuracy, nor does it reflect any intrinsic structure of the function being approximated.

For example, consider the function

$$(x_1, \ldots, x_n) \mapsto \max\{x_1, \ldots, x_n\}.$$

This is a convex piecewise affine function. Using the identity

$$\max(a, b) = a + \sigma(b - a),$$

one can construct this function recursively by composing binary maxima. Such a construction naturally leads to a deep network with linear depth and small width.

On the other hand, representing the same function using only a single hidden layer requires expressing it as a linear combination of many hinge functions of the form $\sigma(\boldsymbol{a}\cdot\boldsymbol{x}+b)$. Although this is theoretically possible, the resulting representation is considerably less direct.

This simple example illustrates an important principle: although shallow networks are universal, deeper networks may represent certain structured functions more naturally and efficiently. In many applications, functions of interest possess a compositional structure. It is therefore natural to consider networks obtained by composing several affine maps and nonlinearities. This leads to the notion of deep (multi-layer) neural networks.

Definition 7.74 (ℓ-layer neural network)**.** Let $\sigma : \mathbb{R} \to \mathbb{R}$. An ℓ-layer neural network with input dimension n and output dimension m is a function of the form

$$\boldsymbol{x} \mapsto W_{\ell+1}\sigma\big(W_\ell\,\sigma(W_{\ell-1}\sigma(\cdots\sigma(W_1\boldsymbol{x} + \boldsymbol{b}_1)\cdots) + \boldsymbol{b}_{\ell-1}) + \boldsymbol{b}_\ell\big),$$

where

$$W_j \in \mathbb{R}^{d_j\times d_{j-1}} \quad j = 1, \cdots, \ell+1, \qquad \boldsymbol{b}_j \in \mathbb{R}^{d_j} \quad j = 1, \ldots, \ell,$$

for suitable intermediate dimensions satisfying

$$d_0 = n, \qquad d_{\ell+1} = m.$$

Remark 7.75. Although a single hidden layer already yields universality, depth may dramatically improve efficiency. There exist functions that can be represented efficiently by deep networks but would require exponentially many neurons if restricted to a single hidden layer. (Precise statements of this phenomenon can be found in the approximation-theoretic literature on depth separation.)

7.7 Kolmogorov–Arnold Representation Theorem

In the previous section we studied the Universal Approximation Theorem, which shows that continuous functions can be approximated by finite linear combinations of ridge functions. Together with the Stone–Weierstrass Theorem, this gives two important examples of approximation results in analysis: one algebraic and one inspired by neural networks.

There is, however, another remarkable way to build multivariable functions from simpler ones. Instead of approximating a function by polynomials or by neural networks, one may ask whether it is possible to represent a multivariable function exactly in terms of continuous functions of a single variable. The Kolmogorov–Arnold representation theorem gives a striking affirmative answer.

Theorem 7.76 (Kolmogorov–Arnold Representation Theorem)**.** *Let $n \geqslant 2$ be an integer. Then there exist positive real numbers*

$$\alpha_1, \cdots, \alpha_n$$

and continuous functions

$$\varphi_1, \cdots, \varphi_{2n+1} : [0,1] \to [0,1]$$

such that for every continuous function $f : [0,1]^n \to \mathbb{R}$ there exists a continuous function

$$h : [0,R] \to \mathbb{R}, \qquad R := \alpha_1 + \cdots + \alpha_n,$$

for which

$$f(x_1, \cdots, x_n) = \sum_{k=1}^{2n+1} h\left(\sum_{m=1}^{n} \alpha_m \varphi_k(x_m)\right) \qquad \forall\, (x_1, \cdots, x_n) \in [0,1]^n. \tag{7.7.1}$$

Remark 7.77. The original formulation of the Kolmogorov–Arnold representation theorem is often written in the form

$$f(\boldsymbol{x}) = \sum_{q=1}^{2n+1} \Phi_q\left(\sum_{p=1}^{n} \psi_{q,p}(x_p)\right),$$

with $2n+1$ *outer functions* Φ_q and $n(2n+1)$ *inner functions* $\psi_{q,p}$. The version presented here is a commonly used modern formulation, in which one works with a single outer function h, together with fixed coefficients $\alpha_1, \cdots, \alpha_n$ and inner functions $\varphi_1, \ldots, \varphi_{2n+1}$.

Remark 7.78. At first sight, Theorem 7.76 may seem rather surprising. However, some familiar multivariable functions already have this form. For example, if

$$f(x_1, \ldots, x_n) = \big(\alpha_1 x_1 + \cdots + \alpha_n x_n\big)^2,$$

then f can be written as

$$f(\boldsymbol{x}) = \sum_{k=1}^{2n+1} h\left(\sum_{m=1}^{n} \alpha_m \varphi_k(x_m)\right)$$

by taking

$$\varphi_1(t) = t, \qquad \varphi_2 = \cdots = \varphi_{2n+1} \equiv 0, \qquad h(t) = t^2.$$

The content of the theorem is that, after choosing suitable functions

$$\varphi_1, \ldots, \varphi_{2n+1},$$

every continuous function on $[0,1]^n$ admits such a representation.

We now turn to the proof of the Kolmogorov–Arnold Representation Theorem. Before giving the proof itself, we introduce some notation and auxiliary definitions. In the following discussion, we fix

$$\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n),$$

where $\alpha_1, \cdots, \alpha_n$ are chosen to be the square roots of n distinct prime numbers. In particular, each α_m is positive.

Given

$$\boldsymbol{\varphi} = (\varphi_1, \cdots, \varphi_{2n+1}) \in \mathscr{C}([0,1]; [0,1]^{2n+1})$$

and a function $h : [0, R] \to \mathbb{R}$, we define

$$(S_{\boldsymbol{\varphi}} h)(\boldsymbol{x}) := \sum_{k=1}^{2n+1} h\left(\sum_{m=1}^{n} \alpha_m \varphi_k(x_m)\right), \qquad \boldsymbol{x} = (x_1, \cdots, x_n) \in [0,1]^n.$$

This is well-defined, because for every $k = 1, \ldots, 2n+1$ and every $\boldsymbol{x} \in [0,1]^n$,

$$0 \leqslant \sum_{m=1}^{n} \alpha_m \varphi_k(x_m) \leqslant \alpha_1 + \cdots + \alpha_n = R.$$

Thus Theorem 7.76 asserts that one can choose a map $\boldsymbol{\varphi}$ such that every $f \in \mathscr{C}([0,1]^n; \mathbb{R})$ is of the form

$$f = S_{\boldsymbol{\varphi}} h$$

for some continuous function $h : [0, R] \to \mathbb{R}$.

Definition 7.79. Let $0 < \lambda < 1$, and let $f : [0,1]^n \to \mathbb{R}$ be bounded. We say that

$$\boldsymbol{\varphi} \in \mathscr{C}([0,1]; [0,1]^{2n+1})$$

is a *λ-prekolmogorov map for f* if either $f \equiv 0$, or else $f \not\equiv 0$ and there exists a continuous function

$$h : [0, R] \to \mathbb{R}$$

such that

$$\|f - S_{\boldsymbol{\varphi}} h\|_\infty < \lambda \|f\|_\infty \qquad \text{and} \qquad \|h\|_\infty \leqslant \|f\|_\infty.$$

We denote by $\mathscr{P}\mathcal{K}_\lambda(f)$ the set of all λ-prekolmogorov maps for f. In particular,

$$\mathscr{P}\mathcal{K}_\lambda(0) = \mathscr{C}([0,1];[0,1]^{2n+1}).$$

Remark 7.80. If $f \in \mathscr{C}([0,1]^n;\mathbb{R})$ and $\boldsymbol{\varphi} \in \mathscr{P}\mathcal{K}_\lambda(f)$, then there exists

$$h \in \mathscr{C}([0,R];\mathbb{R})$$

such that

$$\|f - S_{\boldsymbol{\varphi}} h\|_\infty \leqslant \lambda \|f\|_\infty \qquad \text{and} \qquad \|h\|_\infty \leqslant \|f\|_\infty.$$

Indeed, if $f \not\equiv 0$, this follows immediately from the definition; if $f \equiv 0$, we may simply take $h \equiv 0$.

Remark 7.81. The point of this definition is that a λ-prekolmogorov map does not represent f exactly, but it provides a uniform approximation whose error is at most a fixed proportion λ of $\|f\|_\infty$. Since $\lambda < 1$, such an approximation can be iterated.

The proof below combines a density argument in a suitable function space with an iteration based on uniform convergence. To carry out the density argument, we shall use one general result about complete metric spaces, namely the Baire Category Theorem. A guided proof was given earlier in Problem 3.20, so we only recall the statement here.

Theorem 7.82 (Baire Category Theorem)**.** *Let (X,d) be a complete metric space. If $U_1, U_2, \dots$ are dense open subsets of X, then*

$$\bigcap_{n=1}^{\infty} U_n$$

is dense in X.

The next lemma is the main technical step in the proof. Since its proof is somewhat lengthy, we first explain how it leads to the representation theorem and return to the proof afterward.

Lemma 7.83. *For every $f \in \mathscr{C}([0,1]^n;\mathbb{R})$, the set $\mathscr{P}\mathcal{K}_{\frac{4n+1}{4(n+1)}}(f)$ is open and dense in the metric space*

$$\big(\mathscr{C}([0,1];[0,1]^{2n+1}), d_\infty\big),$$

where for $\boldsymbol{\varphi} = (\phi_1, \dots, \phi_{2n+1})$, $\boldsymbol{\psi} = (\psi_1, \dots, \psi_{2n+1}) \in \mathscr{C}([0,1];[0,1]^{2n+1})$,

$$d_\infty(\boldsymbol{\varphi}, \boldsymbol{\psi}) := \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_\infty = \max_{1 \leqslant k \leqslant 2n+1} \|\phi_k - \psi_k\|_\infty.$$

Lemma 7.84. *There exists $\boldsymbol{\varphi} \in \mathscr{C}([0,1];[0,1]^{2n+1})$ such that*

$$\boldsymbol{\varphi} \in \mathscr{P}\mathcal{K}_{\frac{4n+1}{4(n+1)}}(g)$$

for every polynomial $g : [0,1]^n \to \mathbb{R}$ with rational coefficients.

Proof. Let G be the set of all polynomials on $[0,1]^n$ with rational coefficients. Since G is countable, we may write

$$G = \{g_1, g_2, g_3, \dots\}.$$

By Lemma 7.83, each set $\mathscr{P}\mathcal{K}_{\frac{4n+1}{4(n+1)}}(g_k)$ is open and dense in $\big(\mathscr{C}([0,1];[0,1]^{2n+1}), d_\infty\big)$. Since $\big(\mathscr{C}([0,1];[0,1]^{2n+1}), d_\infty\big)$ is a complete metric space, the Baire Category Theorem implies that

$$\bigcap_{k=1}^{\infty} \mathscr{P}\mathcal{K}_{\frac{4n+1}{4(n+1)}}(g_k)$$

is dense, and in particular nonempty. Hence there exists $\boldsymbol{\varphi} \in \mathscr{C}([0,1];[0,1]^{2n+1})$ such that $\boldsymbol{\varphi} \in \mathscr{P}\mathcal{K}_{\frac{4n+1}{4(n+1)}}(g)$ for every polynomial $g : [0,1]^n \to \mathbb{R}$ with rational coefficients. □

Theorem 7.85. *There exists*

$$\boldsymbol{\varphi} \in \mathscr{C}([0,1];[0,1]^{2n+1})$$

such that

$$\boldsymbol{\varphi} \in \mathscr{P}\mathcal{K}_{\frac{2n+1}{2(n+1)}}(f) \qquad \text{for every } f \in \mathscr{C}([0,1]^n;\mathbb{R}).$$

Equivalently, there exists

$$\boldsymbol{\varphi} \in \mathscr{C}([0,1];[0,1]^{2n+1})$$

such that for every continuous function $f : [0,1]^n \to \mathbb{R}$ there exists

$$h \in \mathscr{C}([0,R];\mathbb{R})$$

satisfying

$$\|f - S_{\boldsymbol{\varphi}} h\|_\infty < \frac{2n+1}{2(n+1)} \|f\|_\infty \qquad \text{and} \qquad \|h\|_\infty \leqslant \|f\|_\infty$$

whenever $f \not\equiv 0$.

Proof. Let $\boldsymbol{\varphi}$ be given by Lemma 7.84, and let

$$f \in \mathscr{C}([0,1]^n; \mathbb{R})$$

be arbitrary. We may assume that $f \not\equiv 0$.

By the Stone–Weierstrass Theorem, there exists a polynomial $g : [0,1]^n \to \mathbb{R}$ with rational coefficients such that

$$\left\| g - \frac{8n+7}{8(n+1)} f \right\|_\infty < \frac{1}{8(n+1)} \|f\|_\infty.$$

Consequently, $g \not\equiv 0$,

$$\|g - f\|_\infty \leqslant \left\| g - \frac{8n+7}{8(n+1)} f \right\|_\infty + \left\| \frac{8n+7}{8(n+1)} f - f \right\|_\infty < \frac{1}{4(n+1)} \|f\|_\infty,$$

and

$$\|g\|_\infty \leqslant \left\| g - \frac{8n+7}{8(n+1)} f \right\|_\infty + \left\| \frac{8n+7}{8(n+1)} f \right\|_\infty < \|f\|_\infty.$$

Since $\boldsymbol{\varphi} \in \mathscr{PK}_{\frac{4n+1}{4(n+1)}}(g)$, there exists $h \in \mathscr{C}([0,R]; \mathbb{R})$ such that

$$\|g - S_{\boldsymbol{\varphi}} h\|_\infty < \frac{4n+1}{4(n+1)} \|g\|_\infty \qquad \text{and} \qquad \|h\|_\infty \leqslant \|g\|_\infty.$$

Hence

$$\begin{aligned} \|f - S_{\boldsymbol{\varphi}} h\|_\infty &\leqslant \|f - g\|_\infty + \|g - S_{\boldsymbol{\varphi}} h\|_\infty < \frac{1}{4(n+1)} \|f\|_\infty + \frac{4n+1}{4(n+1)} \|f\|_\infty \\ &= \frac{2n+1}{2(n+1)} \|f\|_\infty. \end{aligned}$$

Therefore $\boldsymbol{\varphi} \in \mathscr{PK}_{\frac{2n+1}{2(n+1)}}(f)$. □

Proof of Theorem 7.76. Let $\boldsymbol{\varphi}$ be given by Theorem 7.85. We shall construct a sequence of continuous functions

$$h_0, h_1, h_2, \dots \in \mathscr{C}([0,R]; \mathbb{R})$$

such that

$$f = \sum_{\ell=0}^{\infty} S_{\boldsymbol{\varphi}} h_\ell$$

uniformly on $[0,1]^n$.

Set $h_0 := 0$, $f_1 := f$. Assume that for some $k \geqslant 1$ we have already constructed

$$h_0, h_1, \dots, h_{k-1} \in \mathscr{C}([0,R];\mathbb{R}),$$

and define

$$f_k := f - \sum_{\ell=0}^{k-1} S_{\boldsymbol{\varphi}} h_\ell.$$

Since $f_k \in \mathscr{C}([0,1]^n;\mathbb{R})$, Theorem 7.85 yields a function

$$h_k \in \mathscr{C}([0,R];\mathbb{R})$$

such that

$$\|f_k - S_{\boldsymbol{\varphi}} h_k\|_\infty \leqslant \frac{2n+1}{2(n+1)}\|f_k\|_\infty \qquad \text{and} \qquad \|h_k\|_\infty \leqslant \|f_k\|_\infty.$$

Let $\lambda := \dfrac{2n+1}{2(n+1)}$. Since $f_{k+1} = f_k - S_{\boldsymbol{\varphi}} h_k$ for all $k \geqslant 1$, we obtain

$$\|f_{k+1}\|_\infty \leqslant \lambda \|f_k\|_\infty \qquad \forall\, k \geqslant 1.$$

By induction,

$$\|f_{k+1}\|_\infty \leqslant \lambda^k \|f_1\|_\infty = \lambda^k \|f\|_\infty \qquad \forall\, k \geqslant 1.$$

Equivalently,

$$\left\| f - \sum_{\ell=0}^{k} S_{\boldsymbol{\varphi}} h_\ell \right\|_\infty = \|f_{k+1}\|_\infty \leqslant \lambda^k \|f\|_\infty \qquad \forall\, k \geqslant 1. \tag{7.7.2}$$

Moreover,

$$\|h_k\|_\infty \leqslant \|f_k\|_\infty \leqslant \lambda^{k-1}\|f\|_\infty \qquad \forall\, k \geqslant 1.$$

Since $0 < \lambda < 1$, the Weierstrass M-test shows that the series $\sum\limits_{\ell=0}^{\infty} h_\ell$ converges uniformly on $[0,R]$ to a continuous function

$$h : [0,R] \to \mathbb{R}.$$

Now fix $\boldsymbol{x} \in [0,1]^n$. Then

$$\begin{aligned}\left|(S_{\boldsymbol{\varphi}} h)(\boldsymbol{x}) - \sum_{\ell=0}^{N} (S_{\boldsymbol{\varphi}} h_\ell)(\boldsymbol{x})\right| &= \left| \sum_{k=1}^{2n+1} \left[h\left(\sum_{m=1}^{n} \alpha_m \varphi_k(x_m) \right) - \sum_{\ell=0}^{N} h_\ell \left(\sum_{m=1}^{n} \alpha_m \varphi_k(x_m) \right) \right] \right| \\ &\leqslant \sum_{k=1}^{2n+1} \sum_{\ell=N+1}^{\infty} \|h_\ell\|_\infty.\end{aligned}$$

The right-hand side tends to 0 as $N \to \infty$, independently of $\boldsymbol{x}$. Therefore,

$$\lim_{N\to\infty} \left\| S_{\boldsymbol{\varphi}} h - \sum_{\ell=0}^{N} S_{\boldsymbol{\varphi}} h_\ell \right\|_\infty = 0. \tag{7.7.3}$$

Finally, (7.7.2) and (7.7.3) together imply that

$$f = S_{\boldsymbol{\varphi}} h.$$

This completes the proof. ▫

Finally, we turn to the proof of Lemma 7.83.

Proof of "openness" in Lemma 7.83. Let

$$\mu := \frac{4n+1}{4(n+1)}.$$

If $f \equiv 0$, then $\mathscr{PK}_\mu(f) = \mathscr{C}([0,1];[0,1]^{2n+1})$, so there is nothing to prove. Thus we may assume that $f \not\equiv 0$.

Let $\boldsymbol{\psi} \in \mathscr{PK}_\mu(f)$. Then there exists $h : [0,R] \to \mathbb{R}$ such that

$$\|f - S_{\boldsymbol{\psi}} h\|_\infty < \mu \|f\|_\infty \qquad \text{and} \qquad \|h\|_\infty \leqslant \|f\|_\infty.$$

Set

$$\eta := \frac{\mu\|f\|_\infty - \|f - S_{\boldsymbol{\psi}} h\|_\infty}{2n+1} > 0.$$

By uniform continuity of h on $[0,R]$, there exists $\delta > 0$ such that

$$|h(s) - h(t)| < \eta \qquad \text{whenever} \quad |s-t| < \delta,\ \ s,t \in [0,R].$$

We claim that

$$B(\boldsymbol{\psi}, \delta/R) \subseteq \mathscr{PK}_\mu(f).$$

Indeed, let $\boldsymbol{\varphi} \in B(\boldsymbol{\psi}, \delta/R)$. Then for every $\boldsymbol{x} = (x_1, \ldots, x_n) \in [0,1]^n$ and every $k = 1, \ldots, 2n+1$,

$$\left| \sum_{m=1}^{n} \alpha_m \psi_k(x_m) - \sum_{m=1}^{n} \alpha_m \varphi_k(x_m) \right| \leqslant \sum_{m=1}^{n} \alpha_m \|\psi_k - \varphi_k\|_\infty < \sum_{m=1}^{n} \alpha_m \cdot \frac{\delta}{R} = \delta.$$

Hence

$$\left| h\left(\sum_{m=1}^{n} \alpha_m \psi_k(x_m) \right) - h\left(\sum_{m=1}^{n} \alpha_m \varphi_k(x_m) \right) \right| < \eta.$$

Summing over $k = 1, \ldots, 2n+1$, we get

$$\|S_{\psi} h - S_{\varphi} h\|_\infty < (2n+1)\eta.$$

Therefore,

$$\|f - S_{\varphi} h\|_\infty \leqslant \|f - S_{\psi} h\|_\infty + \|S_{\psi} h - S_{\varphi} h\|_\infty < \mu \|f\|_\infty.$$

Since also $\|h\|_\infty \leqslant \|f\|_\infty$, it follows that $\boldsymbol{\varphi} \in \mathscr{PK}_\mu(f)$. This proves that $\mathscr{PK}_\mu(f)$ is open. ▫

The proof of "denseness" in Lemma 7.83 requires some preparations. For each $k = 1, \ldots, 2n+1$, each $N \in \mathbb{N}$ (typically with $N \gg 1$), and each $j \in \mathbb{Z}$, define

$$I_{j,k}(N) := \left[\frac{(2n+1)j + k}{N}, \frac{(2n+1)j + k + 2n}{N}\right] \cap [0,1].$$

Let $\mathbb{Z}_k(N)$ denote the set of all integers j such that $I_{j,k}(N)$ is nonempty; that is,

$$\mathbb{Z}_k(N) = \left\{ j \in \mathbb{Z} \,\middle|\, I_{j,k}(N) \neq \varnothing \right\}.$$

From the definition of $I_{j,k}(N)$, one easily checks that

$$\mathbb{Z}_k(N) = \left[-1, \frac{N-k}{2n+1}\right] \cap \mathbb{Z}.$$

For $i_1, \ldots, i_n \in \mathbb{Z}_k(N)$, we further define

$$S_{i_1, \cdots, i_n; k}(N) := I_{i_1,k}(N) \times I_{i_2,k}(N) \times \cdots \times I_{i_n,k}(N).$$

The next proposition records the basic covering property of these cubes.

Proposition 7.86. *For each fixed positive integer N, the family of cubes*

$$S_{i_1, \cdots, i_n; k}(N), \qquad k = 1, \ldots, 2n+1, \quad i_1, \ldots, i_n \in \mathbb{Z}_k(N),$$

covers $[0,1]^n$ in such a way that every point of $[0,1]^n$ belongs to cubes corresponding to at least $n+1$ distinct values of k.

Proof. Fix $\boldsymbol{x} = (x_1, \ldots, x_n) \in [0,1]^n$. For each $r = 1, \ldots, n$, let

$$E_r := \left\{ k \in \{1, \ldots, 2n+1\} \,\middle|\, \text{there exists } j \in \mathbb{Z}_k(N) \text{ such that } x_r \in I_{j,k}(N) \right\}.$$

We first show that

$$|E_r| \geqslant 2n \qquad \text{for every } r = 1, \ldots, n.$$

Fix $r \in \{1, \dots, n\}$. For each $k \in \{1, \dots, 2n+1\}$, the condition

$$x_r \in I_{j,k}(N)$$

is equivalent to

$$\frac{(2n+1)j + k}{N} \leqslant x_r \leqslant \frac{(2n+1)j + k + 2n}{N} \qquad \text{and} \qquad x_r \in [0,1]$$

or equivalently,

$$Nx_r - 2n \leqslant (2n+1)j + k \leqslant Nx_r, \qquad x_r \in [0,1]$$

Thus, for a fixed k and $x_r \in [0,1]$, such an integer j exists if and only if the interval

$$[Nx_r - 2n - k,\ Nx_r - k]$$

contains a multiple of $2n+1$.

Now this interval has length exactly $2n$, whereas consecutive multiples of $2n+1$ are spaced by $2n+1$. Therefore the interval can fail to contain a multiple of $2n+1$ for at most one value of $k \in \{1, \dots, 2n+1\}$. Hence

$$|E_r| \geqslant 2n.$$

Since $E_r \subseteq \{1, \dots, 2n+1\}$ and each complement E_r^c has cardinality at most 1, we obtain

$$\left| \bigcup_{r=1}^{n} E_r^c \right| \leqslant n.$$

Therefore

$$\left| \bigcap_{r=1}^{n} E_r \right| = (2n+1) - \left| \bigcup_{r=1}^{n} E_r^c \right| \geqslant (2n+1) - n = n+1.$$

So there exist at least $n+1$ distinct values of k such that

$$k \in \bigcap_{r=1}^{n} E_r.$$

Fix such a k. For each $r = 1, \dots, n$, since $k \in E_r$, there exists $i_r \in \mathbb{Z}_k(N)$ such that

$$x_r \in I_{i_r,k}(N).$$

Hence

$$\boldsymbol{x} = (x_1, \dots, x_n) \in I_{i_1,k}(N) \times \cdots \times I_{i_n,k}(N) = S_{i_1,\cdots,i_n;k}(N).$$

Since this happens for at least $n+1$ distinct values of k, the proof is complete. $\square$

Definition 7.87. A function $\varphi \in \mathscr{C}([0,1];[0,1])$ is said to rationally separate a number of pairwise disjoint closed intervals on the line if φ has constant rational pairwise different values on the intervals.

The next proposition shows that a given map can be perturbed slightly so that its values on the relevant intervals become rational and mutually distinct.

Proposition 7.88. *Let ε be a positive number.*

(a) *Let $\psi \in \mathscr{C}([0,1];[0,1])$, and $k \in \{1, \cdots, 2n+1\}$. Then there is an integer $N_0 = N_0(\psi, \varepsilon) > 0$ such that for any integer $N > N_0$ and for any finite set G of rational numbers there is a function $\varphi \in \mathscr{C}([0,1];[0,1])$ such that*

- $\|\varphi - \psi\|_\infty < \varepsilon$;
- *φ rationally separates the family of intervals $I_{j,k}(N)$ parameterized by $j \in \mathbb{Z}_k(N)$;*
- *$\varphi(I_{j,k}(N)) \notin G$ for any $j \in \mathbb{Z}_k(N)$.*

(b) *Let $\boldsymbol{\psi} \in \mathscr{C}([0,1];[0,1]^{2n+1})$. Then there is an integer $N_0 = N_0(\boldsymbol{\psi}, \varepsilon) > 0$ such that for any integer $N > N_0$ there is $\boldsymbol{\varphi} = (\varphi_1, \cdots, \varphi_{2n+1}) \in \mathscr{C}([0,1];[0,1]^{2n+1})$ such that*

- $\|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_\infty < \varepsilon$;
- *for every $k = 1, \cdots, 2n+1$ the function φ_k rationally separates the family of intervals $I_{j,k}(N)$ parameterized by $j \in \mathbb{Z}_k(N)$;*
- *numbers $\varphi_k(I_{j,k}(N))$ are pairwise different for different $k \in \{1, \cdots, 2n+1\}$ and $j \in \mathbb{Z}_k(N)$.*

Proof. Let ε be a positive number.

(a) Let $\psi \in \mathscr{C}([0,1];[0,1])$ and $k \in \{1, \cdots, 2n+1\}$ be given. Since ψ is uniformly continuous on $[0,1]$, there exists $\delta > 0$ such that

$$\big|\psi(x) - \psi(y)\big| < \frac{\varepsilon}{2} \qquad \text{whenever} \quad |x-y| < \delta \text{ and } x, y \in [0,1].$$

Define $N_0 = \left\lceil \dfrac{2n+1}{\delta} \right\rceil$, and let $N > N_0$. Take a piecewise linear function $\varphi \in \mathscr{C}([0,1];[0,1])$ such that

- for each $I_{j,k} := I_{j,k}(N)$, φ is a rational constant on $I_{j,k}$ and
$$\big|\varphi(I_{j,k}) - \psi(\ell_{j,k})\big| < \frac{\varepsilon}{2},$$
where $\ell_{j,k}$ is the left end-point of $I_{j,k}$.
- all values $\varphi(I_{j,k})$ are distinct and $\big\{\varphi(I_{j,k})\big\}_{j\in\mathbb{Z}_k(N)} \subseteq G^{\complement}$.
- φ is linear on the gaps between intervals $I_{j,k}$'s.

Now φ rationally separates the family of intervals $I_{j,k}$ and its values on the intervals are not from G. It remains to show that

$$\big|\varphi(x) - \psi(x)\big| < \varepsilon \qquad \forall\, x \in [0,1]. \tag{7.7.4}$$

(1) Suppose that $x \in I_{j,k}$. Then $\varphi(x) = \varphi(\ell_{j,k})$ and $|x - \ell_{j,k}| \leqslant \dfrac{2n}{N} < \delta$ so we have
$$\big|\varphi(x) - \psi(x)\big| \leqslant \big|\varphi(\ell_{j,k}) - \psi(\ell_{j,k})\big| + \big|\psi(\ell_{j,k}) - \psi(x)\big| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(2) Suppose that $x \in [0,1]\backslash \bigcup\limits_{j\in\mathbb{Z}_k(N)} I_{j,k}$; that is, x lies in the gaps between intervals $I_{j,k}$'s. Suppose that x lies between $I_{j,k}$ and $I_{j+1,k}$. Then $\ell_{j+1,k} - \dfrac{1}{N} < x < \ell_{j+1,k}$, so there exists $\alpha \in (0,1)$ such that
$$x = \ell_{j+1,k} - \frac{\alpha}{N}.$$
Therefore, by the construction of φ we have
$$\varphi(x) = \alpha\varphi\big(\ell_{j+1,k} - \frac{1}{N}\big) + (1-\alpha)\varphi(\ell_{j+1,k}) = \alpha\varphi(\ell_{j,k}) + (1-\alpha)\varphi(\ell_{j+1,k});$$
thus
$$\begin{aligned}
\big|\varphi(x) - \psi(x)\big| &\leqslant \alpha\big|\varphi(\ell_{j,k}) - \psi(x)\big| + (1-\alpha)\big|\varphi(\ell_{j+1,k}) - \psi(x)\big| \\
&\leqslant \alpha\Big[\big|\varphi(\ell_{j,k}) - \psi(\ell_{j,k})\big| + \big|\psi(\ell_{j,k}) - \psi(x)\big|\Big] \\
&\quad + (1-\alpha)\Big[\big|\varphi(\ell_{j+1,k}) - \psi(\ell_{j+1,k})\big| + \big|\psi(\ell_{j+1,k}) - \psi(x)\big|\Big] \\
&< \alpha\Big(\frac{\varepsilon}{2} + \frac{\varepsilon}{2}\Big) + (1-\alpha)\Big(\frac{\varepsilon}{2} + \frac{\varepsilon}{2}\Big) = \varepsilon.
\end{aligned}$$

Therefore, (7.7.4) is established.

(b) Let $\boldsymbol{\psi} = (\psi_1, \cdots, \psi_{2n+1}) \in \mathscr{C}([0,1];[0,1]^{2n+1})$ be given. For each $k \in \{1, \cdots, 2n+1\}$, part (a) provides $N_0(\psi_k, \varepsilon)$, and we define $N_0 = \max\limits_{1 \leqslant k \leqslant 2n+1} N_0(\psi_k, \varepsilon)$. Let $N > N_0$, and as in (a) we write $I_{j,k}$ instead of $I_{j,k}(N)$. Applying part (a) to the case $\psi = \psi_1$, $k = 1$ and $G = \varnothing$, we obtain $\varphi_1 \in \mathscr{C}([0,1];[0,1])$. Then we apply part (a) to the case $\psi = \psi_2$, $k = 2$ and $G = \big\{\varphi_1(I_{j,1}) \,\big|\, j \in \mathbb{Z}_1(N)\big\}$ and we obtain $\varphi_2 \in \mathscr{C}([0,1];[0,1])$. We continue this process: at the m-th step we apply part (a) to the case $\psi = \psi_m$, $k = m$ and $G = \bigcup\limits_{\ell=1}^{m-1} \big\{\varphi_\ell(I_{j,\ell}) \,\big|\, j \in \mathbb{Z}_\ell(N)\big\}$ to obtain $\varphi_m \in \mathscr{C}([0,1];[0,1])$ satisfying

- $\|\varphi_m - \psi_m\|_\infty < \varepsilon$;
- φ_m rationally separates the family of intervals $I_{j,m}(N)$ parameterized by $j \in \mathbb{Z}_m(N)$;
- $\varphi_m(I_{j,m}(N)) \notin G$ for any $j \in \mathbb{Z}_m(N)$.

Then $\boldsymbol{\varphi} = (\varphi_1, \cdots, \varphi_{2n+1})$ satisfies the requirement in part (b). ▫

Proof of "denseness" in Lemma 7.83. If $f \equiv 0$, then

$$\mathscr{P}\mathcal{K}_{\frac{4n+1}{4(n+1)}}(f) = \mathscr{C}([0,1];[0,1]^{2n+1}),$$

so there is nothing to prove. Thus we may assume that $f \in \mathscr{C}([0,1]^n;\mathbb{R})$ is not the zero function.

We establish the denseness result by showing that

$$B(\boldsymbol{\psi}, \varepsilon) \cap \mathscr{P}\mathcal{K}_{\frac{4n+1}{4(n+1)}}(f) \neq \varnothing \qquad \forall\, \boldsymbol{\psi} \in \mathscr{C}([0,1];[0,1]^{2n+1}) \text{ and } \varepsilon > 0.$$

Let $\boldsymbol{\psi} \in \mathscr{C}([0,1];[0,1]^{2n+1})$ and $\varepsilon > 0$ be given. By the uniform continuity of f, there exists $\delta > 0$ such that

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| < \frac{1}{4(n+1)}\|f\|_\infty \qquad \text{whenever} \quad \|\boldsymbol{x} - \boldsymbol{y}\| < \delta,\ \boldsymbol{x}, \boldsymbol{y} \in [0,1]^n.$$

Moreover, Proposition 7.88 (b) provides an $N_0 = N_0(\boldsymbol{\psi}, \varepsilon) > 0$. Choosing $N > N_0$ so large that

$$\frac{2n\sqrt{n}}{N} < \delta,$$

and applying Proposition 7.88 (b) to this N, we obtain a map

$$\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_{2n+1}) \in \mathscr{C}([0,1];[0,1]^{2n+1})$$

such that

$$d_\infty(\boldsymbol{\varphi}, \boldsymbol{\psi}) < \varepsilon, \tag{7.7.5}$$

and, for each $k = 1, \ldots, 2n+1$, the function φ_k is constant on every interval $I_{j,k}(N)$ with $j \in \mathbb{Z}_k(N)$, takes rational values there, and satisfies the separation properties listed in Proposition 7.88 (b).

For notational convenience, for the rest of the proof we write

$$I_{j,k} := I_{j,k}(N), \qquad S_{i_1,\ldots,i_n;k} := S_{i_1,\ldots,i_n;k}(N).$$

Define

$$\widetilde{\varphi}_k(\boldsymbol{x}) := \sum_{m=1}^{n} \alpha_m \varphi_k(x_m), \qquad \boldsymbol{x} = (x_1, \ldots, x_n) \in [0,1]^n,\ 1 \leqslant k \leqslant 2n+1,$$

where we recall that the numbers $\alpha_1, \ldots, \alpha_n$ are the square roots of n distinct prime numbers. Then $\widetilde{\varphi}_k \in \mathscr{C}([0,1]^n; [0,R])$, and we have

$$(S_{\boldsymbol{\varphi}} h)(\boldsymbol{x}) = \sum_{k=1}^{2n+1} h\big(\widetilde{\varphi}_k(\boldsymbol{x})\big). \tag{7.7.6}$$

Moreover, we have the following two consequences.

1. For each fixed $k \in \{1, \ldots, 2n+1\}$, the function $\widetilde{\varphi}_k$ is constant on every cube

   $$S_{i_1,\ldots,i_n;k}.$$

   Moreover, these constant values are pairwise distinct for different cubes with the same k.

   Indeed, suppose that

   $$\widetilde{\varphi}_k(\boldsymbol{x}) = \widetilde{\varphi}_k(\boldsymbol{y}), \qquad \boldsymbol{x} \in S_{i_1,\ldots,i_n;k},\ \boldsymbol{y} \in S_{j_1,\ldots,j_n;k}.$$

   Then

   $$\sum_{m=1}^{n} \alpha_m \varphi_k(x_m) = \sum_{m=1}^{n} \alpha_m \varphi_k(y_m).$$

   Since each $\varphi_k(x_m)$ and $\varphi_k(y_m)$ is rational, and since the numbers

   $$\alpha_1, \ldots, \alpha_n$$

are the square roots of n distinct prime numbers and are linearly independent over $\mathbb{Q}$, it follows that

$$\varphi_k(x_m) = \varphi_k(y_m) \qquad \forall\, m = 1, \dots, n.$$

Because φ_k separates the intervals $I_{j,k}$, we conclude that

$$i_m = j_m \qquad \forall\, m = 1, \dots, n.$$

2. The numbers

$$\widetilde{\varphi}_k\big(S_{i_1,\dots,i_n;k}\big)$$

are pairwise distinct for different $(n+1)$-tuples $(i_1, \dots, i_n, k)$.

Indeed, if

$$\widetilde{\varphi}_k(S_{i_1,\dots,i_n;k}) = \widetilde{\varphi}_\ell(S_{j_1,\dots,j_n;\ell}),$$

then

$$\sum_{m=1}^{n} \alpha_m \varphi_k(I_{i_m,k}) = \sum_{m=1}^{n} \alpha_m \varphi_\ell(I_{j_m,\ell}).$$

Again using the linear independence of $\alpha_1, \dots, \alpha_n$ over $\mathbb{Q}$, we obtain

$$\varphi_k(I_{i_m,k}) = \varphi_\ell(I_{j_m,\ell}) \qquad \forall\, m = 1, \dots, n.$$

By Proposition 7.88, this implies

$$k = \ell \quad \text{and} \quad i_m = j_m \qquad \forall\, m = 1, \dots, n.$$

Because these values are pairwise distinct, the set

$$E := \Big\{ \widetilde{\varphi}_k(S_{i_1,\dots,i_n;k}) \;\Big|\; 1 \leqslant k \leqslant 2n+1,\ i_1, \dots, i_n \in \mathbb{Z}_k \Big\}$$

is a finite subset of $[0, R]$. We now construct a continuous piecewise linear function

$$h : [0, R] \to \mathbb{R}$$

such that

$$h\big(\widetilde{\varphi}_k(S_{i_1,\dots,i_n;k})\big) = \frac{1}{n+1} f(\ell_{i_1,k}, \dots, \ell_{i_n,k})$$

for all $1 \leqslant k \leqslant 2n+1$ and all $i_1, \dots, i_n \in \mathbb{Z}_k$, where $\ell_{j,k}$ denotes the left endpoint of the interval $I_{j,k}$. Moreover, we shall arrange that

$$\|h\|_\infty \leqslant \frac{1}{n+1} \|f\|_\infty.$$

Indeed, for each

$$t = \widetilde{\varphi}_k(S_{i_1,\dots,i_n;k}) \in E,$$

define

$$g(t) := \frac{1}{n+1} f(\ell_{i_1,k}, \dots, \ell_{i_n,k}).$$

Since the points in E are pairwise distinct, this defines a function

$$g : E \to \mathbb{R}.$$

Now list the elements of E in increasing order:

$$0 \leqslant t_1 < \cdots < t_M \leqslant R.$$

We define $h : [0, R] \to \mathbb{R}$ by requiring that

$$h(t_r) = g(t_r) \qquad (r = 1, \dots, M),$$

that h be linear on each interval $[t_r, t_{r+1}]$, and that h be constant on $[0, t_1]$ and on $[t_M, R]$. Then

$$h\big(\widetilde{\varphi}_k(S_{i_1,\dots,i_n;k})\big) = \frac{1}{n+1} f(\ell_{i_1,k}, \dots, \ell_{i_n,k})$$

for every $k = 1, \dots, 2n+1$ and every $i_1, \dots, i_n \in \mathbb{Z}_k$. Moreover, since

$$|g(t)| \leqslant \frac{1}{n+1} \|f\|_\infty \qquad \forall\, t \in E,$$

the piecewise linear construction gives

$$\|h\|_\infty \leqslant \frac{1}{n+1} \|f\|_\infty \leqslant \frac{4n+1}{4(n+1)} \|f\|_\infty. \tag{7.7.7}$$

Next, we show that

$$\big|f(\boldsymbol{z}) - S_{\boldsymbol{\varphi}} h(\boldsymbol{z})\big| < \frac{4n+1}{4(n+1)} \|f\|_\infty \qquad \forall\, \boldsymbol{z} \in [0,1]^n. \tag{7.7.8}$$

Let $\boldsymbol{z} = (z_1, \dots, z_n) \in [0,1]^n$. By Proposition 7.86, there exist at least $n+1$ distinct indices

$$k_1, \dots, k_{n+1} \in \{1, \dots, 2n+1\}$$

such that, for each $s = 1, \dots, n+1$, one can find

$$i_1^{(s)}, \dots, i_n^{(s)} \in \mathbb{Z}_{k_s}(N)$$

with $\boldsymbol{z} \in S_{i_1^{(s)},\dots,i_n^{(s)};k_s}$. Hence

$$z_m \in I_{i_m^{(s)},k_s} \qquad \forall\, m = 1, \dots, n.$$

Therefore

$$\left\| \boldsymbol{z} - \left(\ell_{i_1^{(s)},k_s}, \dots, \ell_{i_n^{(s)},k_s}\right) \right\| \leqslant \frac{2n\sqrt{n}}{N} < \delta,$$

and so

$$\left| f(\boldsymbol{z}) - f\left(\ell_{i_1^{(s)},k_s}, \dots, \ell_{i_n^{(s)},k_s}\right) \right| < \frac{1}{4(n+1)} \|f\|_\infty.$$

Since

$$h\left(\widetilde{\varphi}_{k_s}(\boldsymbol{z})\right) = \frac{1}{n+1} f\left(\ell_{i_1^{(s)},k_s}, \dots, \ell_{i_n^{(s)},k_s}\right),$$

we obtain

$$\left| \frac{1}{n+1} f(\boldsymbol{z}) - h\left(\widetilde{\varphi}_{k_s}(\boldsymbol{z})\right) \right| < \frac{1}{4(n+1)} \|f\|_\infty \qquad \forall\, s = 1, \dots, n+1. \tag{7.7.9}$$

For $k \in \{1, \dots, 2n+1\} \backslash \{k_1, \dots, k_{n+1}\}$, using (7.7.7) we have

$$|h(\widetilde{\varphi}_k(\boldsymbol{z}))| \leqslant \frac{1}{n+1} \|f\|_\infty. \tag{7.7.10}$$

Then (7.7.6), (7.7.9), and (7.7.10) yield

$$\begin{aligned}
|f(\boldsymbol{z}) - S_{\boldsymbol{\varphi}} h(\boldsymbol{z})| &\leqslant \left| f(\boldsymbol{z}) - \sum_{s=1}^{n+1} h(\widetilde{\varphi}_{k_s}(\boldsymbol{z})) \right| + \sum_{k \in \{1,\dots,2n+1\} \backslash \{k_1,\dots,k_{n+1}\}} |h(\widetilde{\varphi}_k(\boldsymbol{z}))| \\
&\leqslant \sum_{s=1}^{n+1} \left| \frac{1}{n+1} f(\boldsymbol{z}) - h(\widetilde{\varphi}_{k_s}(\boldsymbol{z})) \right| + \sum_{k \in \{1,\dots,2n+1\} \backslash \{k_1,\dots,k_{n+1}\}} |h(\widetilde{\varphi}_k(\boldsymbol{z}))| \\
&\leqslant (n+1) \cdot \frac{1}{n+1} \cdot \frac{1}{4(n+1)} \|f\|_\infty + n \cdot \frac{1}{n+1} \|f\|_\infty \\
&= \frac{4n+1}{4(n+1)} \|f\|_\infty.
\end{aligned}$$

Then $\boldsymbol{\varphi} \in B(\boldsymbol{\psi}, \varepsilon) \cap \mathscr{PK}_{\frac{4n+1}{4(n+1)}}(f)$ because of (7.7.5), (7.7.7), and (7.7.8). $\square$

Remark 7.89. The Stone–Weierstrass Theorem, the Universal Approximation Theorem, and the Kolmogorov–Arnold Representation Theorem all describe ways of building complicated functions from simpler ones. Their conclusions, however, are of different types: the first two are approximation theorems, whereas the third gives an exact representation. In addition, the Kolmogorov–Arnold theorem shows that the inner functions can be chosen independently of the target function f.

We conclude this section with a comparison of the Stone–Weierstrass Theorem, the Universal Approximation Theorem, and the Kolmogorov–Arnold Representation Theorem.

	Stone–Weierstrass	Universal Approximation	Kolmogorov–Arnold
Type of result	Density theorem	Density theorem	Exact representation
Setting	Algebra of functions	Neural networks	Superposition of univariate functions
Basic building block	Polynomials (or algebra generators)	Ridge functions $\sigma(\boldsymbol{a} \cdot x + b)$	$h(\sum_{m=1}^{n} \alpha_m \varphi_k(x_m))$
Mechanism	Algebraic generation	Linear superposition	Compositional structure
Approximation v.s. identity	Approximation	Approximation	Exact identity
Dependence on target function	Coefficients depend on f	Parameters depend on f	Only outer functions depend on f
Structural insight	Closed subalgebra dense in $\mathscr{C}(K)$	Ridge functions dense in $\mathscr{C}(K)$	Multivariate functions reduce to univariate compositions

Conceptual Summary.

- Stone–Weierstrass shows that sufficiently rich algebras generate all continuous functions by uniform approximation.
- The Universal Approximation Theorem shows that finite linear combinations of ridge functions are uniformly dense in $\mathscr{C}(K)$.
- The Kolmogorov–Arnold theorem shows that every multivariate continuous function can be written exactly as a finite sum of superpositions of univariate functions.

7.8 Exercises

§ 7.1 Pointwise and Uniform Convergence

Problem 7.1. Let (M, d) and (N, ρ) be metric spaces, $A \subseteq M$, and $f_k : A \to N$ be a sequence of functions such that for some function $f : A \to N$, we have that for all $x \in A$, if

$\{x_k\}_{k=1}^{\infty} \subseteq A$ and $x_k \to x$ as $k \to \infty$, then

$$\lim_{k\to\infty} f_k(x_k) = f(x)\,.$$

Show that

1. $\{f_k\}_{k=1}^{\infty}$ converges pointwise to f.

2. If $\{f_{k_j}\}_{j=1}^{\infty}$ is a subsequence of $\{f_k\}_{k=1}^{\infty}$, and $\{x_j\}_{j=1}^{\infty} \subseteq A$ is a convergent sequence satisfying that $\lim\limits_{j\to\infty} x_j = x$, then

$$\lim_{j\to\infty} f_{k_j}(x_j) = f(x)\,.$$

3. Show that if in addition A is compact and f is continuous on A, then $\{f_k\}_{k=1}^{\infty}$ converges uniformly f on A.

Remark. Using the inequality

$$\rho\big(f_k(x_k), f(x)\big) \leqslant \rho\big(f(x_k), f(x)\big) + \sup_{x\in A} \rho\big(f_k(x), f(x)\big)\,,$$

we find that if $\{f_k\}_{k=1}^{\infty}$ converges uniformly to a continuous function f, then $\lim\limits_{k\to\infty} f_k(x_k) = f(x)$ as long as $\lim\limits_{k\to\infty} x_k = x$. Together with the conclusion in 3, we conclude that

> Let (M,d), (N,ρ) be metric spaces, $K \subseteq M$ be a compact set, $f_k : K \to N$ be a function for each $k \in \mathbb{N}$, and $f : K \to N$ be continuous. The sequence $\{f_k\}_{k=1}$ converges uniformly to f if and only if $\lim\limits_{k\to\infty} f_k(x_k) = f(x)$ whenever sequence $\{x_k\}_{k=1}^{\infty} \subseteq K$ converges to x.

Problem 7.2. Let (M,d) be a metric space, $A \subseteq M$, (N,ρ) be a complete metric space, and $f_k : A \to N$ be a sequence of functions (not necessary continuous) which converges uniformly on A. Suppose that $a \in \mathrm{cl}(A)$ and

$$\lim_{x\to a} f_k(x) = L_k$$

exists for all $k \in \mathbb{N}$. Show that $\{L_k\}_{k=1}^{\infty}$ converges, and

$$\lim_{x\to a}\lim_{k\to\infty} f_k(x) = \lim_{k\to\infty}\lim_{x\to a} f_k(x)\,.$$

Problem 7.3. Prove the Dini theorem:

> Let K be a compact set, and $f_k : K \to \mathbb{R}$ be continuous for all $k \in \mathbb{N}$ such that $\{f_k\}_{k=1}$ converges pointwise to a continuous function $f : K \to \mathbb{R}$. Suppose that $f_k \leqslant f_{k+1}$ for all $k \in \mathbb{N}$. Then $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on K.

Hint: Mimic the proof of showing that $\{c_k\}_{k=1}^{\infty}$ converges to 0 in Lemma 6.64.

Problem 7.4. Let (M, d) and (N, ρ) be metric spaces, $A \subseteq M$, and $f_k : A \to N$ be uniformly continuous functions, and $\{f_k\}_{k=1}^{\infty}$ converges uniformly to $f : A \to N$ on A. Show that f is uniformly continuous on A.

Problem 7.5. Let (M, d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a norm space, $B \subseteq A \subseteq M$, $f_k : A \to \mathcal{V}$ be bounded for each $k \in \mathbb{N}$, and $\{g_n\}_{n=1}^{\infty}$ be the Cesàro mean of $\{f_k\}_{k=1}^{\infty}$; that is, $g_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} f_k$. Show that if $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on B, then $\{g_n\}_{n=1}^{\infty}$ converges uniformly to f on B.

Problem 7.6. Complete the following.

1. Suppose that $f_k, f, g : [0, \infty) \to \mathbb{R}$ are functions such that

   (a) $\forall\, R > 0$, f_k and g are Riemann integrable on $[0, R]$;

   (b) $|f_k(x)| \leqslant g(x)$ for all $k \in \mathbb{N}$ and $x \in [0, \infty)$;

   (c) $\forall\, R > 0$, $\{f_k\}_{k=1}^{\infty}$ converges to f uniformly on $[0, R]$;

   (d) $\displaystyle\int_0^{\infty} g(x)dx \equiv \lim_{R\to\infty}\int_0^R g(x)dx < \infty$.

   Show that $\displaystyle\lim_{k\to\infty}\int_0^{\infty} f_k(x)dx = \int_0^{\infty} f(x)dx$; that is,
   $$\lim_{k\to\infty}\lim_{R\to\infty}\int_0^R f_k(x)dx = \lim_{R\to\infty}\lim_{k\to\infty}\int_0^R f_k(x)dx\,.$$

2. Let $f_k(x)$ be given by $f_k(x) = \begin{cases} 1 & \text{if } k-1 \leqslant x < k\,, \\ 0 & \text{otherwise.}\end{cases}$ Find the (pointwise) limit f of the sequence $\{f_k\}_{k=1}^{\infty}$, and check whether $\displaystyle\lim_{k\to\infty}\int_0^{\infty} f_k(x)dx = \int_0^{\infty} f(x)dx$ or not. Briefly explain why one can or cannot apply 1.

3. Let $f_k : [0, \infty) \to \mathbb{R}$ be given by $f_k(x) = \dfrac{x}{1 + kx^4}$. Find $\lim\limits_{k\to\infty} \displaystyle\int_0^\infty f_k(x)dx$.

§ 7.2 Series of Functions and The Weierstrass M-Test

Problem 7.7. Show that the series

$$\sum_{k=1}^{\infty}(-1)^k \frac{x^2 + k}{k^2}$$

converges uniformly on every bounded interval.

Problem 7.8. Consider the function

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{1 + k^2 x}\,.$$

On what intervals does it converge uniformly? On what intervals does it fail to converge uniformly? Is f continuous wherever the series converges? If f bounded?

Problem 7.9. Determine which of the following real series $\sum\limits_{k=1}^{\infty} g_k$ converge (pointwise or uniformly). Check the continuity of the limit in each case.

1. $g_k(x) = \begin{cases} 0 & \text{if } x \leqslant k\,, \\ (-1)^k & \text{if } x > k\,. \end{cases}$

2. $g_k(x) = \begin{cases} \dfrac{1}{k^2} & \text{if } |x| \leqslant k\,, \\ \dfrac{1}{x^2} & \text{if } |x| > k\,. \end{cases}$

3. $g_k(x) = \dfrac{(-1)^k}{\sqrt{k}} \cos(kx)$ on $\mathbb{R}$.

§ 7.3 Integration and Differentiation of Series

Problem 7.10. In the following series of functions defined on $\mathbb{R}$, find its domain of convergence (classify it into domain of absolute and conditional convergence). If the series is a power series, find its radius of convergence. Also discuss whether the series is uniformly convergent in every compact subsets of its domain of convergence. Determine which series can be differentiated or integrated term by term in its domain of convergence.

(1) $\sum\limits_{k=1}^{\infty} \dfrac{x}{k^\alpha + k^\beta x^2}$, $\alpha \geqslant 0$, $\beta > 0$;

(2) $\sum\limits_{k=1}^{\infty} \frac{1}{2^k}\sqrt{1-x^{2k}}$;

(3) $\sum\limits_{k=1}^{\infty} \frac{1\cdot 3\cdots(2k-1)}{2\cdot 4\cdots(2k)}\big(1+\frac{1}{2}+\cdots+\frac{1}{k}\big)x^{2k}$;

(4) $\sum\limits_{k=1}^{\infty} \frac{(-1)^{k-1}}{k\log(k+1)}x^{k!}$;

(5) $\sum\limits_{k=1}^{\infty} a_k x^k$, where $\{a_k\}_{k=1}^{\infty}$ is defined by the recursive relation $a_k = 3a_{k-1} - 2a_{k-2}$ for $k \geqslant 2$, and $a_0 = 1$, $a_1 = 2$.

Also find the sum of the series in (5).

Problem 7.11. In this problem we investigate the differentiability of a complex power series. This requires a new proof of $\frac{d}{dx}\sum\limits_{k=0}^{\infty} a_k x^k = \sum\limits_{k=1}^{\infty} k a_k x^{k-1}$ instead of making use of Theorem 7.11.

Let $\{a_k\}_{k=0}^{\infty} \subseteq \mathbb{R}$ be a real sequence, and $f(x) = \sum\limits_{k=0}^{\infty} a_k x^k$ be a (real) power series with radius of convergence $R > 0$. Let $s_n(x) = \sum\limits_{k=0}^{n} a_k x^k$ be the n-th partial sum, $R_n(x) = f(x) - s_n(x)$, and $g(x) = \sum\limits_{k=1}^{\infty} k a_k x^{k-1}$. For $x, x_0 \in (-\rho, \rho) \subsetneq (-R, R)$, where $x \neq x_0$, write

$$\frac{f(x)-f(x_0)}{x-x_0} - g(x) = \frac{s_n(x)-s_n(x_0)}{x-x_0} - s_n'(x_0) + \big(s_n'(x_0) - g(x_0)\big) + \frac{R_n(x)-R_n(x_0)}{x-x_0}. \quad (7.8.1)$$

1. Show that
$$\Big|\frac{R_n(x)-R_n(x_0)}{x-x_0}\Big| \leqslant \sum_{k=n+1}^{\infty} k|a_k|\rho^{k-1},$$
and use the inequality above to show that $\lim\limits_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = g(x_0)$.

2. Generalize the conclusion to complex power series; that is, show that if $\{a_k\}_{k=0}^{\infty} \subseteq \mathbb{C}$ and the power series $\sum\limits_{k=0}^{\infty} a_k z^k$ has radius of convergence $R > 0$, then
$$\frac{d}{dz}\sum_{k=0}^{\infty} a_k z^k = \sum_{k=1}^{\infty} k a_k z^{k-1} \qquad \forall\, |z| < R.$$
Assume that you have known $\frac{d}{dz}\sum\limits_{k=0}^{n} a_k z^k = \sum\limits_{k=1}^{n} k a_k z^{k-1}$ for all $n \in \mathbb{N} \cup \{0\}$ (this can be proved using the definition of differentiability of functions with values in normed vector spaces provided in Chapter 5).

Problem 7.12. Suppose that the series $\sum_{n=0}^{\infty} a_n = 0$, and $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for $-1 < x \leqslant 1$. Show that f is continuous at $x = 1$ by complete the following.

1. Write $s_n = a_0 + a_1 + \cdots + a_n$ and $S_n(x) = a_0 + a_1 x + \cdots + a_n x^n$. Show that
$$S_n(x) = (1-x)(s_0 + s_1 x + \cdots + s_{n-1}x^{n-1}) + s_n x^n$$
and $f(x) = (1-x)\sum_{n=0}^{\infty} s_n x^n$.

2. Using the representation of f from above to conclude that $\lim_{x\to 1^-} f(x) = 0$.

3. What if $\sum_{n=0}^{\infty} a_n$ is convergent but not zero?

Problem 7.13. Construct the function $g(x)$ by letting $g(x) = |x|$ if $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ and extending g so that it becomes periodic (with period 1). Define
$$f(x) = \sum_{k=1}^{\infty} \frac{g(4^{k-1}x)}{4^{k-1}}.$$

1. Use the Weierstrass M-test to show that f is continuous on $\mathbb{R}$.

2. Prove that f is differentiable at no point.

(So there exists a continuous which is nowhere differentiable!)

Hint: Google Blancmange function!

§ 7.4 The Space of Continuous Functions

Problem 7.14. Let $\delta : (\mathscr{C}([-1,1];\mathbb{R}), \|\cdot\|_\infty) \to \mathbb{R}$ be defined by $\delta(f) = f(0)$. Show that δ is linear and uniformly continuous.

Problem 7.15. Let (M,d) be a metric space, and $K \subseteq M$ be a compact subset.

1. Show that the set $U = \{f \in \mathscr{C}(K;\mathbb{R}) \,|\, a < f(x) < b \text{ for all } x \in K\}$ is open in $(\mathscr{C}(K;\mathbb{R}), \|\cdot\|_\infty)$ for all $a, b \in \mathbb{R}$.

2. Show that the set $F = \{f \in \mathscr{C}(K;\mathbb{R}) \,|\, a \leqslant f(x) \leqslant b \text{ for all } x \in K\}$ is closed in $(\mathscr{C}(K;\mathbb{R}), \|\cdot\|_\infty)$ for all $a, b \in \mathbb{R}$.

3. Let $A \subseteq M$ be a subset, not necessarily compact. Prove or disprove that the set $B = \{f \in \mathscr{C}_b(A;\mathbb{R}) \,|\, f(x) > 0 \text{ for all } x \in A\}$ is open in $\big(\mathscr{C}_b(A;\mathbb{R}), \|\cdot\|_\infty\big)$.

§ 7.5 The Stone-Weierstrass Theorem

Problem 7.16. Define B to be the set of all even functions in the space $\mathscr{C}([-1,1];\mathbb{R})$; that is, $f \in B$ if and only if f is continuous on $[-1,1]$ and $f(x) = f(-x)$ for all $x \in [-1,1]$. Prove that B is closed but not dense in $\mathscr{C}([-1,1];\mathbb{R})$. Hence show that even polynomials are dense in B, but not in $\mathscr{C}([-1,1];\mathbb{R})$.

Problem 7.17. Let $f : [0,1] \to \mathbb{R}$ be a continuous function.

1. Suppose that
$$\int_0^1 f(x)x^n dx = 0 \qquad \forall\, n \in \mathbb{N} \cup \{0\}\,.$$
Show that $f = 0$ on $[0,1]$.

2. Suppose that for some $m \in \mathbb{N}$,
$$\int_0^1 f(x)x^n dx = 0 \qquad \forall\, n \in \{0, 1, \cdots, m\}\,.$$
Show that $f(x) = 0$ has at least $(m+1)$ distinct real roots around which $f(x)$ change signs.

Problem 7.18. Let $f : [0,1] \to \mathbb{R}$ be continuous. Show that
$$\lim_{n\to\infty} \int_0^1 f(x)\cos(nx)\,dx = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \int_0^1 f(x)\sin(nx)\,dx = 0\,.$$

Problem 7.19. Put $p_0 = 0$ and define
$$p_{k+1}(x) = p_k(x) + \frac{x^2 - p_k^2(x)}{2} \qquad \forall\, k \in \mathbb{N} \cup \{0\}\,.$$
Show that $\{p_k\}_{k=1}^\infty$ converges uniformly to $|x|$ on $[-1,1]$.

Hint: Use the identity
$$|x| - p_{k+1}(x) = \big[|x| - p_k(x)\big]\Big[1 - \frac{|x| + p_k(x)}{2}\Big] \tag{7.8.2}$$
to prove that $0 \leqslant p_k(x) \leqslant p_{k+1}(x) \leqslant |x|$ if $|x| \leqslant 1$, and that
$$|x| - p_k(x) \leqslant |x|\Big(1 - \frac{|x|}{2}\Big)^k < \frac{2}{k+1}$$
if $|x| \leqslant 1$.

Problem 7.20. Let $f : [0,1] \to \mathbb{R}$ be continuous and $\varepsilon > 0$. Prove that there is a simple function g (defined in Example 7.55) such that $\|f - g\|_\infty < \varepsilon$.

Problem 7.21. Suppose that p_n is a sequence of polynomials converging uniformly to f on $[0,1]$ and f is not a polynomial. Prove that the degrees of p_n are not bounded.
Hint: An Nth-degree polynomial p is uniquely determined by its values at $N+1$ points $x_0, \cdots, x_N$ via Lagrange's interpolation formula

$$p(x) = \sum_{k=0}^{N} \pi_k(x) \frac{p(x_k)}{\pi_k(x_k)},$$

where $\pi_k(x) = (x - x_0)(x - x_1) \cdots (x - x_N)/(x - x_k) = \prod\limits_{\substack{1 \leqslant j \leqslant N \\ j \neq k}} (x - x_j)$.

Problem 7.22. Consider the set of all functions on $[0,1]$ of the form

$$h(x) = \sum_{j=1}^{n} a_j e^{b_j x},$$

where a_j, $b_j \in \mathbb{R}$. Is this set dense in $\mathscr{C}([0,1];\mathbb{R})$?

§ 7.6 Universal Approximation Theorem

Problem 7.23 (Density for generalized ridge functions)**.** This problem asks you to generalize Proposition 7.67. Fix an integer d with $1 \leqslant d \leqslant n-1$, and consider functions of the form

$$G(\boldsymbol{x}) = g(A\boldsymbol{x}),$$

where $A \in \mathbb{R}^{d\times n}$ and $g : \mathbb{R}^d \to \mathbb{R}$ is continuous. (When $d = 1$, this reduces to the usual ridge functions $f(\boldsymbol{a} \cdot \boldsymbol{x})$.)

Let Ω be a subset of $\mathbb{R}^{d\times n}$ and define

$$\mathcal{M}(\Omega) := \operatorname{span}\left\{ g(A\boldsymbol{x}) \;\middle|\; A \in \Omega,\ g \in \mathscr{C}(\mathbb{R}^d;\mathbb{R}) \right\}.$$

For each $A \in \Omega$, let $L(A)$ denote the row space of A (equivalently, the column space of $A^\top$), and set

$$L(\Omega) := \bigcup_{A\in\Omega} L(A) \subseteq \mathbb{R}^n.$$

Prove the "$\Leftarrow$" direction of the following statement.

Proposition 7.90. *The linear space $\mathcal{M}(\Omega)$ is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$ in the topology of uniform convergence on compact subsets if and only if the only polynomial in H^n that vanishes identically on $L(\Omega)$ is the zero polynomial.*

The proof of the opposite direction will be given later (see Problem 9.46 in Chapter 9).

Hint. Show that $(\boldsymbol{d}\cdot\boldsymbol{x})^k \in \mathcal{M}(\Omega)$ for every $\boldsymbol{d}\in L(\Omega)$ and every $k\in\mathbb{N}$. Then mimic the argument in Proposition 7.67.

Problem 7.24 (A smooth-activation version of UAT)**.** In this problem you are asked to prove the following weaker version of the Universal Approximation Theorem.

Theorem 7.91. *Let $\sigma\in\mathscr{C}^\infty(\mathbb{R};\mathbb{R})$. Then σ is not a polynomial if and only if $\mathcal{N}_\sigma(n,m)$ is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R}^m)$ in the sense of uniform convergence on compact subsets.*

The direction "$\Leftarrow$" is proved in Theorem 7.72. Your task is to prove "$\Rightarrow$".

Hint. It suffices to show that $\Sigma_1(\sigma)$ is dense in $\mathscr{C}(\mathbb{R};\mathbb{R})$. Use difference quotients to show that for each $k\in\mathbb{N}$,

$$\frac{d^k}{dw^k}\,\sigma(wx+\theta)\in\overline{\Sigma_1(\sigma)}^{\|\cdot\|_\infty}\qquad (w,\theta\in\mathbb{R}),$$

where the closure is taken with respect to uniform convergence on compact sets. By the chain rule, this implies

$$x^k\,\sigma^{(k)}(wx+\theta)\in\overline{\Sigma_1(\sigma)}^{\|\cdot\|_\infty}\qquad (k\in\mathbb{N},\ w,\theta\in\mathbb{R}).$$

In particular,

$$x^k\,\sigma^{(k)}(\theta)\in\overline{\Sigma_1(\sigma)}^{\|\cdot\|_\infty}\qquad (k\in\mathbb{N},\ \theta\in\mathbb{R}).$$

Use this to show that $\overline{\Sigma_1(\sigma)}^{\|\cdot\|_\infty}$ contains all polynomials. Finally, apply the Stone–Weierstrass Theorem.

Problem 7.25 (UAT for a sigmoidal activation)**.** Let $\sigma:\mathbb{R}\to\mathbb{R}$ be a continuous *sigmoidal* function; that is,

$$\lim_{t\to-\infty}\sigma(t)=0,\qquad \lim_{t\to+\infty}\sigma(t)=1.$$

Show that the set $\mathcal{N}_\sigma(n,1)$ is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$.

Background. A classical result of Cybenko shows that such σ yields uniform density of $\mathcal{N}_\sigma(n,1)$ in $\mathscr{C}(K;\mathbb{R})$.

Hints:

1. It suffices to prove the result for $K = [0,1]^n$.

2. Set
$$S := \operatorname{span}\left\{\sigma(\boldsymbol{a}\cdot\boldsymbol{x}+b) \,\middle|\, \boldsymbol{a}\in\mathbb{R}^n,\ b\in\mathbb{R}\right\} \subset \mathscr{C}([0,1]^n).$$
Assume that S is not dense. By the Hahn–Banach theorem, there exists a nonzero continuous linear functional L on $\mathscr{C}([0,1]^n)$ such that $L(S) = \{0\}$.

3. By the Riesz representation theorem, there exists a finite signed Borel measure μ on $[0,1]^n$ such that
$$L(h) = \int_{[0,1]^n} h(\boldsymbol{x})\,d\mu(\boldsymbol{x}) \quad \text{for all } h \in \mathscr{C}([0,1]^n;\mathbb{R}).$$
Hence
$$\int_{[0,1]^n} \sigma(\boldsymbol{a}\cdot\boldsymbol{x}+b)\,d\mu(\boldsymbol{x}) = 0 \quad \text{for all } \boldsymbol{a}\in\mathbb{R}^n,\ b\in\mathbb{R}.$$

4. Use the sigmoidal property: for fixed $\boldsymbol{a} \neq 0$, the functions $\sigma(t(\boldsymbol{a}\cdot\boldsymbol{x}+b))$ converge pointwise to the indicator function of the half-space $\{\boldsymbol{a}\cdot\boldsymbol{x}+b>0\}$ as $t\to+\infty$. Apply dominated convergence (note that σ is bounded) to conclude that
$$\mu\big(\{\boldsymbol{a}\cdot\boldsymbol{x}+b>0\}\big) = 0 \quad \text{for all } \boldsymbol{a}\in\mathbb{R}^n,\ b\in\mathbb{R}.$$

5. Deduce that μ must vanish identically. This contradicts the assumption that $L \neq 0$. Hence S is dense in $\mathscr{C}([0,1]^n)$, and the desired density follows.

§ 7.7 Kolmogorov-Arnold Representation Theorem

Problem 7.26 (The remainder of the original Kolmogorov–Arnold construction)**.** In the original Kolmogorov–Arnold approach, the most delicate part is the construction of certain numbers
$$\lambda^{pq}_{k,i} \qquad (1 \leqslant p \leqslant n,\ 1 \leqslant q \leqslant 2n+1,\ 1 \leqslant i \leqslant m_k = (9n)^k+1,\ k \geqslant 0)$$
together with a sequence $\{\varepsilon_k\}_{k=0}^{\infty}$ of positive numbers. In this problem, we assume that this part has already been carried out, and we show that the remainder of the proof may then be completed by standard arguments involving uniform convergence and continuity.

Assume that for each $k \geqslant 0$, $1 \leqslant p \leqslant n$, and $1 \leqslant q \leqslant 2n+1$, there are numbers
$$\lambda^{pq}_{k,1} < \lambda^{pq}_{k,2} < \cdots < \lambda^{pq}_{k,m_k}$$
and positive numbers ε_k satisfying the following properties:

(1) For each fixed k, p, q,

$$\lambda^{pq}_{k,i+1} \leqslant \lambda^{pq}_{k,i} + 2^{-k} \qquad (i = 1, \ldots, m_k - 1).$$

(2) If $A^q_{k,i} \cap A^q_{k-1,j} \neq \varnothing$, then

$$[\lambda^{pq}_{k,i},\ \lambda^{pq}_{k,i} + \varepsilon_k] \subseteq [\lambda^{pq}_{k-1,j},\ \lambda^{pq}_{k-1,j} + \varepsilon_{k-1}].$$

(3) For each fixed k and q, the intervals

$$\Delta^q_{k,i_1,\ldots,i_n} := \left[\sum_{p=1}^{n} \lambda^{pq}_{k,i_p},\ \sum_{p=1}^{n} \lambda^{pq}_{k,i_p} + n\varepsilon_k\right]$$

are pairwise disjoint as $(i_1, \ldots, i_n)$ ranges over $\{1, \ldots, m_k\}^n$.

(4) The series $\sum\limits_{k=0}^{\infty} \varepsilon_k$ converges.

For $1 \leqslant q \leqslant 2n+1$, $k \in \mathbb{N}$, and $1 \leqslant i \leqslant m_k$, let

$$A^q_{k,i} = \left[\frac{1}{(9n)^k}\left(i - 1 - \frac{q}{3n}\right), \frac{1}{(9n)^k}\left(i - \frac{1}{3n} - \frac{q}{3n}\right)\right],$$

and for each multi-index $(i_1, \ldots, i_n) \in \{1, \ldots, m_k\}^n$, define

$$S^q_{k,i_1,\ldots,i_n} := A^q_{k,i_1} \times \cdots \times A^q_{k,i_n}.$$

Thus the index set $\{1, \ldots, m_k\}$ plays the same role as the sets $\mathbb{Z}_k(N)$ used in the main text: it simply enumerates those intervals $A^q_{k,i}$ that meet $[0,1]$.

For each fixed k, p, q, define a function

$$\psi^{pq}_k : [0,1] \to \mathbb{R}$$

as follows: on each interval $A^q_{k,i}$, let ψ^{pq}_k be equal to the constant $\lambda^{pq}_{k,i}$, and on each connected component of

$$[0,1] \backslash \bigcup_{i=1}^{m_k} A^q_{k,i},$$

define ψ^{pq}_k by linear interpolation. Then $\psi^{pq}_k \in \mathscr{C}([0,1]; \mathbb{R})$ for every k, p, q.

Complete the following.

1. Show that for each fixed k, the family of cubes

$$\left\{S^q_{k,i_1,\ldots,i_n} :\ 1 \leqslant q \leqslant 2n+1,\ (i_1,\ldots,i_n) \in \{1,\ldots,m_k\}^n\right\}$$

covers the unit cube $[0,1]^n$, and in fact every point of $[0,1]^n$ belongs to at least $n+1$ such cubes.

2. Show that

$$\|\psi^{pq}_k - \psi^{pq}_{k-1}\|_\infty \leqslant \varepsilon_{k-1} \qquad \forall\, k \geqslant 1.$$

Hint: Use property (2), and consider separately the points lying in some interval $A^q_{k,i}$ and the points lying in the gaps, where both functions are defined by linear interpolation.

3. Deduce that for each fixed p, q, the sequence $\{\psi^{pq}_k\}_{k=0}^\infty$ is uniformly Cauchy on $[0,1]$, and hence converges uniformly to a continuous function

$$\psi^{pq} \in \mathscr{C}([0,1];\mathbb{R}).$$

4. For each fixed q, define

$$\Psi^q(\boldsymbol{x}) := \sum_{p=1}^n \psi^{pq}(x_p), \qquad \boldsymbol{x} = (x_1,\ldots,x_n) \in [0,1]^n.$$

Show that for every $k \geqslant 0$, every multi-index $(i_1,\ldots,i_n)$, and every $\boldsymbol{x} \in S^q_{k,i_1,\ldots,i_n}$, one has

$$\Psi^q(\boldsymbol{x}) \in \Delta^q_{k,i_1,\ldots,i_n}.$$

Hint: Fix $x_p \in A^q_{k,i_p}$. Show that for every $j \geqslant k$,

$$\psi^{pq}_j(x_p) \in [\lambda^{pq}_{k,i_p},\ \lambda^{pq}_{k,i_p} + \varepsilon_k],$$

using property (2), and then pass to the limit $j \to \infty$.

5. Deduce from property (3) that, for each fixed k and q, distinct cubes

$$S^q_{k,i_1,\ldots,i_n}$$

are mapped by Ψ^q into pairwise disjoint intervals. More precisely, if

$$(i_1,\ldots,i_n) \neq (j_1,\ldots,j_n),$$

then

$$\Psi^q\big(S^q_{k,i_1,\dots,i_n}\big) \subseteq \Delta^q_{k,i_1,\dots,i_n}, \qquad \Psi^q\big(S^q_{k,j_1,\dots,j_n}\big) \subseteq \Delta^q_{k,j_1,\dots,j_n},$$

and these two intervals are disjoint.

6. Let

$$f_0 := f.$$

For each $r \geqslant 1$, suppose that continuous functions

$$\chi^q_{r-1} : \mathbb{R} \to \mathbb{R}, \qquad 1 \leqslant q \leqslant 2n+1,$$

have already been constructed, and define the residual function

$$f_{r-1}(\boldsymbol{x}) := f(\boldsymbol{x}) - \sum_{q=1}^{2n+1} \chi^q_{r-1}(\Psi^q(\boldsymbol{x})), \qquad \boldsymbol{x} \in [0,1]^n.$$

Explain why, by the uniform continuity of f_{r-1}, one can choose an integer k_r so large that the oscillation of f_{r-1} on each cube $S^q_{k_r,i_1,\dots,i_n}$ is bounded by

$$\sup_{\boldsymbol{x},\boldsymbol{y} \in S^q_{k_r,i_1,\dots,i_n}} \big|f_{r-1}(\boldsymbol{x}) - f_{r-1}(\boldsymbol{y})\big| \leqslant \frac{1}{2n+2} \|f_{r-1}\|_\infty. \tag{7.8.3}$$

7. Fix $r \geqslant 1$ and $q \in \{1, \dots, 2n+1\}$. For each multi-index $(i_1, \dots, i_n)$, choose a point

$$\xi^q_{k_r,i_1,\dots,i_n} \in S^q_{k_r,i_1,\dots,i_n}$$

(for instance, the center of the cube). Define a function η^q_r on the union of the intervals

$$\Delta^q_{k_r,i_1,\dots,i_n}$$

by

$$\eta^q_r(y) = \frac{1}{n+1} f_{r-1}\big(\xi^q_{k_r,i_1,\dots,i_n}\big), \qquad y \in \Delta^q_{k_r,i_1,\dots,i_n}.$$

Use property (3) to show that this definition is well-defined. Then extend η^q_r continuously to the whole real line by linear interpolation on the gaps between these intervals.

8. Define
$$\chi_r^q := \chi_{r-1}^q + \eta_r^q, \qquad 1 \leqslant q \leqslant 2n+1.$$
Show that
$$\|\eta_r^q\|_\infty \leqslant \frac{1}{n+1}\|f_{r-1}\|_\infty,$$
and hence
$$\|\chi_r^q - \chi_{r-1}^q\|_\infty \leqslant \frac{1}{n+1}\|f_{r-1}\|_\infty. \tag{7.8.4}$$

9. Show, using (7.8.3), that the new residual
$$f_r(\boldsymbol{x}) := f(\boldsymbol{x}) - \sum_{q=1}^{2n+1} \chi_r^q(\Psi^q(\boldsymbol{x}))$$
satisfies
$$\|f_r\|_\infty \leqslant \frac{2n+1}{2(n+1)}\|f_{r-1}\|_\infty.$$
Hint: Note that
$$f_r(x_1,\ldots,x_n) = f_{r-1}(x_1,\ldots,x_n) - \sum_q \left[\chi_r^q\Big(\sum_p \psi^{pq}(x_p)\Big) - \chi_{r-1}^q\Big(\sum_p \psi^{pq}(x_p)\Big)\right].$$
Split the sum $\sum_q$ as $\sum' + \sum''$, where $\sum'$ extends over $n+1$ values of q such that $\boldsymbol{x} = (x_1,\ldots,x_n)$ belongs to one of the cubes $S^q_{k_r,i_1,\ldots,i_n}$, whose existence is guaranteed by Step 1, and $\sum''$ extends over the remaining n values of q. Then use (7.8.3) and (7.8.4).

10. Show that for each fixed q, the series
$$\sum_{r=1}^\infty (\chi_r^q - \chi_{r-1}^q)$$
converges uniformly on every bounded interval. Deduce that χ_r^q converges uniformly on every bounded interval to a continuous function χ^q, and conclude that
$$f(\boldsymbol{x}) = \sum_{q=1}^{2n+1} \chi^q(\Psi^q(\boldsymbol{x})) \qquad \forall\, \boldsymbol{x} \in [0,1]^n.$$
Hint: First show that $\Psi^q([0,1]^n)$ is bounded.

11. Conclude that, once the numbers $\lambda_{k,i}^{pq}$ and the sequence $\{\varepsilon_k\}$ with properties (1)–(4) have been constructed, the rest of the original Kolmogorov–Arnold proof follows from standard arguments involving uniform convergence, continuity, and geometric decay of the residuals.

§ 7.8 Chapter Review

Problem 7.27 (**True or False).** Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. Let $f_n : [a, b] \to \mathbb{R}$ be an uniformly convergent sequence of continuous functions. Then the sequence of the indefinite integrals $g_n(x)$ defined by

$$g_n(x) = \int_a^x f_n(t)\, dt \qquad \forall\, x \in [a, b]$$

converges uniformly to a continuously differentiable function.

2. Let $f_n : [0, 1] \to \mathbb{R}$ be a equi-continuous sequence of functions such that the sequence $\big\{f_n(\frac{1}{2})\big\}_{n=1}^{\infty}$ is bounded in $\mathbb{R}$. Then $\{f_n\}_{n=1}^{\infty}$ contains a convergent subsequence.

Chapter 8

Fourier Series

讓我們回顧一下之前已經有的一些結論。在 § 7.5 中我們學到了 Stone-Weierstrass 定理，它告訴我們定義在 $[0,1]$ 上的連續函數 f 可以用多項式（例如 Bernstein 多項式）去逼近（在均勻收斂的意義下），而我們也注意到 Bernstein 多項式，在取不同次數 n 的多項式做逼近時，每一個單項式 x^k 前面的係數都跟 n 和 k 有關。但是從定理 7.23 中我們又發現，對某些擁有很好性質的函數 f（叫做解析函數 Analytic functions），即使取不同次數 n 的多項式做逼近時，每個單項式 x^k 前面的係數可以取成只跟函數 f 的 k 次導數有關（跟 n 無關）。這給了我們一個很粗略的概念，知道想用多項式去逼近連續函數時，多項式的係數有些時候會跟多項次的次數有關，有時則無關。

在這一章中，我們在前四節特別關注在週期為 2π 的連續函數。由 Corollary 7.65 我們知道這樣的函數可用形如

$$p_n(x) = \frac{c_0^{(n)}}{2} + \sum_{k=1}^{n} (c_k^{(n)} \cos kx + s_k^{(n)} \sin kx)$$

的三角多項式 (trigonometric polynomials) 所逼近（在均勻收斂的意義下），其中上標 (n) 代表的是係數可能與用來逼近的三角多項式的次數 n 有關係。跟前一段所述的經驗類似，在數學理論上我們想知道下面問題的答案：

1. 什麼樣的函數，可以用係數與逼近次數無關的三角多項式去逼近。對這樣的函數，三角多項式要怎麼挑？

2. 對於實在沒辦法用係數與逼近次數無關的三角多項式去逼近的連續週期函數，有什麼好的方法逼近？而上面所挑出來的那個係數跟逼近次數無關的三角多項式，在次數接近無窮大時出了什麼問題？

上述的問題解決之後，我們用變數變換，也可以得到對於週期為 $2L$ 的函數的相關理論。

另外，由於在進行的過程中，我們發現我們所關心用來逼近連續函數的三角多項式（叫富氏級數），其係數的定法只要求函數可積分即可，因此，一個自然衍生的問題則是：對不連續（但可積分）的函數來說，有沒有什麼收斂理論可以說明？這個部份的研究則是第四、五節的主要重點。在第六節中，我們則提供了一個快速傅利葉變換 (FFT) 的演算法可供電腦去計算富氏級數（的係數）。

8.1 Basic Properties of the Fourier Series

Let $f \in \mathscr{C}(\mathbb{T})$ be given. We first assume that the trigonometric polynomials used to approximate f can be chosen in such a way that the coefficients does not depend on the degree of approximation; that is, $c_k^{(n)} = c_k$ and $s_k^{(n)} = s_k$. In this case, if $p_n \to f$ uniformy on $[-\pi, \pi]$, by Theorem 7.17 we must have

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} p_n(x)\cos kx\,dx = \int_{-\pi}^{\pi} f(x)\cos kx\,dx \qquad \forall\, k \in \{0, 1, \cdots, n\}$$

and

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} p_n(x)\sin kx\,dx = \int_{-\pi}^{\pi} f(x)\sin kx\,dx \qquad \forall\, k \in \{1, \cdots, n\}.$$

Since

$$\int_{-\pi}^{\pi} \cos kx\cos \ell x\,dx = \int_{-\pi}^{\pi} \sin kx\sin \ell x\,dx = \pi\delta_{k\ell} \qquad \forall\, k, \ell \in \mathbb{N}$$

and

$$\int_{-\pi}^{\pi} \sin kx\cos \ell x\,dx = 0 \qquad \forall\, k \in \mathbb{N}, \ell \in \mathbb{N} \cup \{0\},$$

we find that

$$c_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx \qquad \text{and} \qquad s_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\,dx. \tag{8.1.1}$$

This induces the following

Definition 8.1. For a Riemann integrable function $f : [-\pi, \pi] \to \mathbb{R}$, the ***Fourier series*** of f, denoted by $s(f, \cdot)$, is given by

$$s(f, x) = \frac{c_0}{2} + \sum_{k=1}^{\infty} (c_k \cos kx + s_k \sin kx)$$

whenever the sum makes sense, where sequences $\{c_k\}_{k=0}^{\infty}$ and $\{s_k\}_{k=1}^{\infty}$ given by (8.1.1) are called the ***Fourier coefficients*** associated with f. The n-th partial sum of the Fourier series to f, denoted by $s_n(f,\cdot)$, is given by

$$s_n(f,x) = \frac{c_0}{2} + \sum_{k=1}^{n} (c_k \cos kx + s_k \sin kx).$$

We note that for the Fourier series $s(f,x)$ to be defined, f is not necessary continuous. Our goal is to establish the convergence of Fourier series in various senses.

Remark 8.2. Because of the Euler identity $e^{i\theta} = \cos\theta + i\sin\theta$, we can write

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)(e^{iky} + e^{-iky})dy \quad \text{and} \quad s_k = \frac{1}{2\pi i}\int_{-\pi}^{\pi} f(y)(e^{iky} - e^{-iky})dy$$

thus

$$\begin{aligned} s_n(f,x) &= \frac{c_0}{2} + \sum_{k=1}^{n}\big(c_k\frac{e^{ikx}+e^{-ikx}}{2} + s_k\frac{e^{ikx}-e^{-ikx}}{2i}\big) \\ &= \frac{1}{2}\Big[c_0 + \sum_{k=1}^{n}\big((c_k - is_k)e^{ikx} + (c_k + is_k)e^{-ikx}\big)\Big] \\ &= \frac{1}{2}\Big[c_0 + \sum_{k=1}^{n}(c_k - is_k)e^{ikx} + \sum_{k=-n}^{-1}(c_{-k} + is_{-k})e^{ikx}\Big] \\ &= \frac{1}{2}\Big[c_0 + \frac{1}{\pi}\sum_{k=1}^{n}\int_{-\pi}^{\pi} f(y)e^{-iky}dye^{ikx} + \frac{1}{\pi}\sum_{k=-n}^{-1}\int_{-\pi}^{\pi} f(y)e^{-iky}dye^{ikx}\Big]. \end{aligned}$$

Define $\widehat{f}_k = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} f(y)e^{-iky}\,dy$. Then $\widehat{f}_k = \dfrac{c_{|k|} + is_{|k|}}{2}$ (here we treat $s_0 = 0$), and

$$s_n(f,x) = \sum_{k=-n}^{n}\widehat{f}_k e^{ikx}.$$

The sequence $\{\widehat{f}_k\}_{k=-\infty}^{\infty}$ is also called the Fourier coefficients associated with f, and one can write the Foruier series of f as $\sum\limits_{k=-\infty}^{\infty}\widehat{f}_k e^{ikx}$.

Remark 8.3. Given a continuous function g with period $2L$ (or a function g which is Riemann integrable on $[-L,L]$), let $f(x) = g\big(\frac{Lx}{\pi}\big)$. Then f is a continuous function with period 2π (or f is a Riemann integrable function on $[-\pi, \pi]$), and the Fourier series of f is given by

$$s(f, x) = \frac{c_0}{2} + \sum_{k=1}^{n} (c_k \cos kx + s_k \sin kx),$$

where c_k and s_k are given by (8.1.1). Now, define the Fourier series of g by $s(g, x) = s\big(f, \frac{\pi x}{L}\big)$. Then the Fourier series of g is given by

$$s(g, x) = \frac{c_0}{2} + \sum_{k=1}^{\infty} \Big(c_k \cos \frac{k\pi x}{L} + s_k \sin \frac{k\pi x}{L}\Big),$$

where $\{c_k\}_{k=0}^{\infty}$ and $\{s_k\}_{k=1}^{\infty}$ is also called the Fourier coefficients associated with g and are given by

$$c_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx = \frac{1}{\pi} \int_{-\pi}^{\pi} g\Big(\frac{Lx}{\pi}\Big) \cos kx \, dx = \frac{1}{L} \int_{-L}^{L} g(x) \cos \frac{k\pi x}{L} \, dx$$

and similarly, $s_k = \dfrac{1}{L} \displaystyle\int_{-L}^{L} g(x) \sin \frac{k\pi x}{L} \, dx$. Similar to Remark 8.2, the Fourier series of g can also be written as

$$\sum_{k=-\infty}^{\infty} \widehat{g}_k e^{\frac{i\pi kx}{L}},$$

where $\widehat{g}_k = \dfrac{1}{2L} \displaystyle\int_{-L}^{L} g(y) e^{\frac{-i\pi ky}{L}} \, dy$.

Example 8.4. Consider the periodic function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x & \text{if } 0 \leqslant x \leqslant \pi, \\ -x & \text{if } -\pi < x < 0, \end{cases}$$

and $f(x + 2\pi) = f(x)$ for all $x \in \mathbb{R}$. To find the Fourier representation of f, we compute the Fourier coefficients by

$$s_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx \, dx = \frac{1}{\pi} \Big(\int_0^{\pi} x \sin kx \, dx - \int_{-\pi}^{0} x \sin kx \, dx \Big) = 0$$

and

$$c_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx = \frac{1}{\pi} \Big(\int_0^{\pi} x \cos kx \, dx - \int_{-\pi}^{0} x \cos kx \, dx \Big) = \frac{2}{\pi} \int_0^{\pi} x \cos kx \, dx.$$

If $k = 0$, then $c_0 = \dfrac{2}{\pi} \displaystyle\int_0^{\pi} x \, dx = \pi$, while if $k \in \mathbb{N}$,

$$c_k = \frac{2}{\pi} \Big(\frac{x \sin kx}{k} \Big|_0^{\pi} - \int_0^{\pi} \frac{\sin kx}{k} \, dx \Big) = \frac{2}{\pi} \frac{\cos kx}{k^2} \Big|_0^{\pi} = \frac{2((-1)^k - 1)}{\pi k^2}.$$

Therefore, $c_{2k} = 0$ and $c_{2k-1} = -\dfrac{4}{\pi(2k-1)^2}$ for all $k \in \mathbb{N}$. Therefore, the Fourier series of f is given by

$$s(f,x) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=1}^{\infty} \frac{\cos(2k-1)x}{(2k-1)^2}.$$

Example 8.5. Consider the periodic function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } -\dfrac{\pi}{2} \leqslant x \leqslant \dfrac{\pi}{2}, \\ 0 & \text{if } -\pi \leqslant x < -\dfrac{\pi}{2} \text{ or } \dfrac{\pi}{2} < x \leqslant \pi, \end{cases}$$

and $f(x+2\pi) = f(x)$ for all $x \in \mathbb{R}$. We compute the Fourier coefficients of f and find that $s_k = 0$ for all $k \in \mathbb{N}$ and $c_0 = 1$, as well as

$$c_k = \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos kx\,dx = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \cos kx\,dx = \frac{2\sin\frac{k\pi}{2}}{\pi k}.$$

Therefore, $c_{2k} = 0$ and $c_{2k-1} = \dfrac{2(-1)^{k+1}}{\pi(2k-1)}$ for all $k \in \mathbb{N}$; thus the Fourier series of f is given by

$$s(f,x) = \frac{1}{2} - \frac{2}{\pi}\sum_{k=1}^{\infty} \frac{(-1)^k}{2k-1}\cos(2k-1)x.$$

Example 8.6. Consider the periodic function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = x \qquad \text{if} \quad -\pi < x \leqslant \pi$$

and $f(x+2\pi) = f(x)$ for all $x \in \mathbb{R}$. Then the Fourier coefficients of f are computed as follows: $c_k = 0$ for all $k \in \mathbb{N} \cup \{0\}$ since f is (more or less) an odd function, and

$$\begin{aligned} s_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x\sin kx\,dx = \frac{2}{\pi}\int_0^{\pi} x\sin kx\,dx = \frac{2}{\pi}\Big(-\frac{x\cos kx}{k}\Big|_0^{\pi} + \int_0^{\pi}\frac{\cos kx}{k}\,dx\Big) \\ &= \frac{2(-1)^{k+1}}{k}. \end{aligned}$$

Therefore, the Fourier series of f is given by

$$s(f,x) = 2\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\sin kx.$$

8.2 Uniform Convergence of the Fourier Series

Before proceeding, we note that Remark 8.2 implies that

$$s_n(f,x) = \sum_{k=-n}^{n} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{ik(x-y)}\,dy = \int_{-\pi}^{\pi} f(y)\Big(\frac{1}{2\pi}\sum_{k=-n}^{n} e^{ik(x-y)}\Big)\,dy\,.$$

Define $D_n(x) = \dfrac{1}{2\pi}\sum\limits_{k=-n}^{n} e^{ikx}$. Then D_n is 2π-periodic, and

$$s_n(f,x) = \int_{-\pi}^{\pi} f(y)D_n(x-y)\,dy\,.$$

For 2π-periodic Riemann integrable functions f and g, we define the convolution of f and g on the circle by

$$(f \star g)(x) = \int_{-\pi}^{\pi} f(y)g(x-y)\,dy\,.$$

Then $s_n(f,x) = (D_n \star f)(x)$.

Note that $D_n(0) = \dfrac{2n+1}{2\pi}$, and if $e^{ix} \neq 1$,

$$D_n(x) = \frac{1}{2\pi}\frac{e^{-inx}\big[e^{i(2n+1)x}-1\big]}{e^{ix}-1} = \frac{1}{2\pi}\frac{e^{i(n+1/2)x}-e^{-i(n+1/2)x}}{e^{ix/2}-e^{-ix/2}} = \frac{\sin(n+\frac{1}{2})x}{2\pi\sin\frac{x}{2}}$$

so that we have the following

Definition 8.7. The function $D_n : \mathbb{R} \to \mathbb{R}$ defined by

$$D_n(x) = \begin{cases} \dfrac{\sin(n+\frac{1}{2})x}{2\pi\sin\frac{x}{2}} & \text{if } x \notin \{2k\pi \,|\, k \in \mathbb{Z}\}\,, \\ \dfrac{2n+1}{2\pi} & \text{if } x \in \{2k\pi \,|\, k \in \mathbb{Z}\}\,, \end{cases} \tag{8.2.1}$$

is called the ***Dirichlet kernel***.

By the fact that $D_n(x) = \dfrac{1}{2\pi}\sum\limits_{k=-n}^{n} e^{ikx}$, we immediately conclude the following

Lemma 8.8. *For each $n \in \mathbb{N}$ and $x \in \mathbb{R}$,* $\displaystyle\int_{-\pi}^{\pi} D_n(x-y)\,dy = 1$.

In the following, we first consider the uniform convergence of the Fourier series of 2π-periodic continuously differentiable functions.

Definition 8.9. The normed vector space $\big(\mathscr{C}^1(\mathbb{T}), \|\cdot\|_{\mathscr{C}^1(\mathbb{T})}\big)$ is a vector space over $\mathbb{R}$ consisting of all 2π-periodic real-valued continuously differentiable functions and is equipped with a norm

$$\|f\|_{\mathscr{C}^1(\mathbb{T})} = \|f\|_\infty + \|f'\|_\infty = \max_{x\in\mathbb{R}} \big|f(x)\big| + \max_{x\in\mathbb{R}} \big|f'(x)\big| \qquad \forall\, f \in \mathscr{C}^1(\mathbb{T}).$$

Theorem 8.10. *For any $f \in \mathscr{C}^1(\mathbb{T})$,the Fourier series of f converges uniformly to f on $\mathbb{R}$; that is, the sequence $\{s_n(f,\cdot)\}_{n=1}^\infty$ converges uniformly to f on $\mathbb{R}$.*

Proof. By Lemma 8.8, we find that for all $x \in \mathbb{R}$,

$$\begin{aligned} s_n(f,x) - f(x) = (D_n \star f - f)(x) &= \int_{-\pi}^{\pi} D_n(x-y)\big(f(y) - f(x)\big)\,dy \\ &= \int_{-\pi}^{\pi} D_n(y)\big(f(x-y) - f(x)\big)\,dy\,. \end{aligned}$$

We break the integral into two parts: one is the integral on $|y| \leqslant \delta$ and the other is the integral on $\delta < |y| \leqslant \pi$. Since $f \in \mathscr{C}^1(\mathbb{T})$,

$$|f(x-y) - f(x)| \leqslant \|f'\|_\infty |y|\,;$$

thus by the fact that $\dfrac{x}{\sin x} \leqslant \dfrac{\pi}{2}$ for $0 < x < \dfrac{\pi}{2}$, we obtain that

$$\begin{aligned} &\Big|\int_{|y|\leqslant\delta} D_n(y)\big(f(x-y) - f(x)\big)\,dy\Big| \\ &\qquad \leqslant \int_{-\delta}^{\delta} \frac{\big|f(x-y) - f(x)\big|}{2\pi\big|\sin\frac{y}{2}\big|}\,dy \leqslant \frac{\|f'\|_\infty}{2\pi}\int_{-\delta}^{\delta} \frac{y}{\sin\frac{y}{2}}\,dy \leqslant \|f'\|_\infty \delta\,. \end{aligned} \tag{8.2.2}$$

Now we take care of the integral on $\delta < |y| \leqslant \pi$ by first looking at the integral on $\delta < y < \pi$. Integrating by parts,

$$\begin{aligned} \int_\delta^\pi D_n(y)\big(f(x-y) - f(y)\big)\,dy &= \frac{1}{2\pi}\int_\delta^\pi \sin\big(n+\frac{1}{2}\big)y\,\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}\,dy \\ = -\frac{1}{2\pi}\frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}\Big|_{y=\delta}^{y=\pi} &+ \frac{1}{2\pi}\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{d}{dy}\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}\,dy\,. \end{aligned}$$

For the first term on the right-hand side,

$$\Big|\frac{1}{2\pi}\frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}\Big|_{y=\delta}^{y=\pi}\Big| \leqslant \frac{2\|f\|_\infty}{2\pi n\sin\frac{\delta}{2}} \leqslant \frac{\|f\|_\infty}{n\sin\frac{\delta}{2}} \qquad \forall\, x \in \mathbb{R}\,.$$

For the second term on the right-hand side,

$$\begin{aligned}
&\Big|\frac{1}{2\pi}\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{d}{dy}\frac{f(x-y)-f(x)}{\sin\frac{y}{2}}dy\Big| \\
&\qquad \leqslant \frac{1}{2\pi}\bigg[\Big|\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{f'(x-y)}{\sin\frac{y}{2}}dy\Big| + \Big|\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{\cos\frac{y}{2}\big(f(x-y)-f(x)\big)}{2\sin^2\frac{y}{2}}dy\Big|\bigg] \\
&\qquad \leqslant \frac{1}{2\pi}\Big[\|f'\|_\infty\frac{\pi-\delta}{\big(n+\frac{1}{2}\big)\sin\frac{\delta}{2}} + \|f\|_\infty\frac{\pi-\delta}{\big(n+\frac{1}{2}\big)\sin^2\frac{\delta}{2}}\Big] \leqslant \frac{\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}}.
\end{aligned}$$

Similarly,

$$\Big|\int_{-\pi}^{-\delta} D_n(y)\big(f(x-y)-f(x)\big)dy\Big| \leqslant \frac{\|f\|_\infty}{n\sin\frac{\delta}{2}} + \frac{\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}};$$

thus for all $x \in \mathbb{R}$,

$$\begin{aligned}
\big|s_n(f,x)-f(x)\big| &\leqslant \Big|\Big(\int_{-\delta}^{\delta} + \int_\delta^\pi + \int_{-\pi}^{-\delta}\Big)D_n(y)\big(f(x-y)-f(x)\big)\,dy\Big| \\
&\leqslant \|f'\|_\infty\delta + \frac{2\|f\|_\infty}{n\sin\frac{\delta}{2}} + \frac{2\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}} \leqslant \|f'\|_\infty\delta + \frac{4\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}}.
\end{aligned}$$

Let $\varepsilon > 0$ be given. Choose a fixed $\delta > 0$ such that $\|f'\|_\infty\delta < \dfrac{\varepsilon}{2}$. For this fixed δ, choose $N > 0$ such that

$$\frac{4\|f\|_{\mathscr{C}^1(\mathbb{T})}}{N\sin^2\frac{\delta}{2}} < \frac{\varepsilon}{2}.$$

Then if $n \geqslant N$ and $x \in \mathbb{R}$, we have

$$\big|s_n(f,x)-f(x)\big| < \frac{\varepsilon}{2} + \frac{4\|f\|_{\mathscr{C}^1(\mathbb{T})}}{n\sin^2\frac{\delta}{2}} \leqslant \frac{\varepsilon}{2} + \frac{4\|f\|_{\mathscr{C}^1(\mathbb{T})}}{N\sin^2\frac{\delta}{2}} < \varepsilon. \qquad \square$$

Next we consider the convergence of the Fourier series of less regular functions. The functions of which we prove the convergence of the Fourier series belong to the so-called Hölder class continuous functions.

Definition 8.11. A function $f \in \mathscr{C}(\mathbb{T})$ is said to be ***Hölder continuous with exponent*** $\alpha \in (0,1]$, denoted by $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, if $\displaystyle\sup_{\substack{x,y\in\mathbb{R}\\ x\neq y}} \frac{\big|f(x)-f(y)\big|}{|x-y|^\alpha} < \infty$. Let $\|\cdot\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})}$ be defined by

$$\|f\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})} = \sup_{x\in\mathbb{T}}\big|f(x)\big| + \sup_{\substack{x,y\in\mathbb{R}\\ x\neq y}}\frac{\big|f(x)-f(y)\big|}{|x-y|^\alpha}.$$

Then $\|\cdot\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})}$ is a norm on $\mathscr{C}^{0,\alpha}(\mathbb{T})$, and

$$\mathscr{C}^{0,\alpha}(\mathbb{T}) = \big\{f \in \mathscr{C}(\mathbb{T}) \,\big|\, \|f\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})} < \infty\big\}.$$

In particular, when $\alpha = 1$, a function in $\mathscr{C}^{0,1}(\mathbb{T})$ is said to be Lipschitz continuous on $\mathbb{T}$; thus $\mathscr{C}^{0,1}(\mathbb{T})$ consists of Lipschitz continuous functions on $\mathbb{T}$.

The uniform convergence of $s_n(f,\cdot)$ to f for $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ with $\alpha \in (0,1)$ requires a lot more work. The idea is to estimate $\|f - s_n(f,\cdot)\|_\infty$ in terms of the quantity $\inf\limits_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_\infty$. Since $s_n(f,\cdot) \in \mathscr{P}_n(\mathbb{T})$, it is obvious that

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_\infty \leqslant \|f - s_n(f,\cdot)\|_\infty .$$

The goal is to show the inverse inequality

$$\|f - s_n(f,\cdot)\|_\infty \leqslant C_n \inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_\infty \tag{8.2.3}$$

for some constant C_n, and pick a suitable $p \in \mathscr{P}_n(\mathbb{T})$ which gives a good upper bound for $\|f - s_n(f,\cdot)\|_\infty$. The inverse inequality is established via the following

Proposition 8.12. *The Dirichlet kernel D_n satisfies that for all $n \in \mathbb{N}$,*

$$\int_{-\pi}^{\pi} \big|D_n(x)\big| dx \leqslant 2 + \log n . \tag{8.2.4}$$

Proof. The validity of (8.2.4) for the case $n = 1$ is left to the reader, and we provide the proof for the case $n \geqslant 2$ here. Recall that $D_n(x) = \dfrac{\sin(n+\frac{1}{2})x}{2\pi \sin\frac{x}{2}}$ if $x \in (0,\pi]$. Therefore,

$$\int_{-\pi}^{\pi} \big|D_n(x)\big| dx = 2\int_0^{\pi} \big|D_n(x)\big| dx = \int_0^{\frac{1}{n}} 2\big|D_n(x)\big| dx + \int_{\frac{1}{n}}^{\pi} \Big|\frac{\sin(n+\frac{1}{2})x}{\pi \sin\frac{x}{2}}\Big| dx .$$

Since $|D_n(x)| \leqslant \dfrac{2n+1}{2\pi}$ for all $0 < x \leqslant \dfrac{1}{n}$, the first integral can be estimated by

$$\int_0^{\frac{1}{n}} 2\big|D_n(x)\big| dx \leqslant \frac{1}{\pi}\frac{2n+1}{n} . \tag{8.2.5}$$

Since $\dfrac{2x}{\pi} \leqslant \sin x$ for $0 \leqslant x \leqslant \dfrac{\pi}{2}$, the second integral can be estimated by

$$\int_{\frac{1}{n}}^{\pi} \Big|\frac{\sin(n+\frac{1}{2})x}{\pi \sin\frac{x}{2}}\Big| dx \leqslant \int_{\frac{1}{n}}^{\pi} \frac{1}{x}\, dx = \log \pi + \log n . \tag{8.2.6}$$

We then conclude (8.2.4) from (8.2.5) and (8.2.6) by noting that $\log \pi + \dfrac{2n+1}{n\pi} \leqslant 2$ for all $n \geqslant 2$. $\square$

Remark 8.13. A more subtle estimate can be done to show that

$$\int_{-\pi}^{\pi} \big|D_n(x)\big| dx \geqslant c_1 + c_2 \log n \qquad \forall\, n \in \mathbb{N}$$

for some positive constants c_1 and c_2. Therefore, the integral of $|D_n|$ on $[-\pi, \pi]$ blows up as $n \to \infty$.

With the help of Proposition 8.12, we are able to prove the inverse inequality (8.2.3). The following theorem is a direct consequence of Proposition 8.12.

Theorem 8.14. *Let $f \in \mathscr{C}(\mathbb{T})$; that is, f is a continuous function with period 2π. Then*

$$\big\|f - s_n(f, \cdot)\big\|_\infty \leqslant (3 + \log n) \inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty . \tag{8.2.7}$$

Proof. For $n \in \mathbb{N}$ and $x \in \mathbb{T}$,

$$\big|s_n(f, x)\big| \leqslant \int_{-\pi}^{\pi} \big|D_n(y)\big|\big|f(x-y)\big| dy \leqslant (2 + \log n)\|f\|_\infty .$$

Given $\varepsilon > 0$, let $p \in \mathscr{P}_n(\mathbb{T})$ such that

$$\|f - p\|_\infty \leqslant \inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty + \varepsilon.$$

Then by the fact that $s_n(p, x) = p(x)$ if $p \in \mathscr{P}_n(\mathbb{T})$, we obtain that

$$\begin{aligned}
\big\|f - s_n(f, \cdot)\big\|_\infty &\leqslant \big\|f - p\big\|_\infty + \big\|p - s_n(f, \cdot)\big\|_\infty \leqslant \big\|f - p\big\|_\infty + \big\|s_n(f - p, \cdot)\big\|_\infty \\
&\leqslant \big\|f - p\big\|_\infty + (2 + \log n)\|f - p\|_\infty \\
&\leqslant (3 + \log n)\Big[\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty + \varepsilon\Big],
\end{aligned}$$

and (8.2.7) is obtained by passing to the limit as $\varepsilon \to 0$. □

Having established Theorem 8.14, the study of the uniform convergence of $s_n(f, \cdot)$ to f then amounts to the study of the quantity $\inf\limits_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty$. The estimate of $\inf\limits_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty$ for $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, where $\alpha \in (0, 1)$, is more difficult, and requires a clever choice of p. We begin with the following

Lemma 8.15. *If f is a continuous function on $[a, b]$, then for all $\delta_1, \delta_2 > 0$,*

$$\sup_{|x-y| \leqslant \delta_1} \big|f(x) - f(y)\big| \leqslant \Big(1 + \frac{\delta_1}{\delta_2}\Big) \sup_{|x-y| \leqslant \delta_2} \big|f(x) - f(y)\big| .$$

The proof of Lemma 8.15 is not very difficult, and is left to the readers.

Now we are in position to prove the theorem due to D. Jackson.

Theorem 8.16 (Jackson)**.** *There exists a constant $C > 0$ such that*

$$\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty \leqslant C \sup_{|x-y| \leqslant \frac{1}{n}} |f(x) - f(y)| \qquad \forall\, f \in \mathscr{C}(\mathbb{T}).$$

Proof. Let $p(x) = 1 + c_1 \cos x + \cdots + c_n \cos nx$ be a positive trigonometric function of degree n with coefficients $\{c_i\}_{i=1}^n$ determined later. Define an operator K on $\mathscr{C}(\mathbb{T})$ by

$$(\mathrm{K}f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} p(y) f(x-y)\, dy.$$

Then $\mathrm{K}f \in \mathscr{P}_n(\mathbb{T})$. Lemma 8.15 then implies

$$\begin{aligned}
\big|(\mathrm{K}f)(x) - f(x)\big| &\leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} p(y) \big|f(x-y) - f(x)\big| dy \\
&\leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} p(y) \big(1 + n|y|\big) \sup_{|x-y| \leqslant \frac{1}{n}} \big|f(x) - f(y)\big| dy \\
&= \Big[1 + \frac{n}{2\pi} \int_{-\pi}^{\pi} |y| p(y)\, dy\Big] \sup_{|x-y| \leqslant \frac{1}{n}} \big|f(x) - f(y)\big|.
\end{aligned}$$

Since $y^2 \leqslant \dfrac{\pi^2}{2}(1 - \cos y)$ for $y \in [-\pi, \pi]$, by the Cauchy-Schwarz inequality (Corollary 2.27) we find that

$$\begin{aligned}
\frac{1}{2\pi} \int_{-\pi}^{\pi} |y| p(y)\, dy &\leqslant \Big[\frac{1}{2\pi} \int_{-\pi}^{\pi} y^2 p(y)\, dy\Big]^{\frac{1}{2}} \Big[\frac{1}{2\pi} \int_{-\pi}^{\pi} p(y)\, dy\Big]^{\frac{1}{2}} \\
&\leqslant \Big[\frac{\pi}{4} \int_{-\pi}^{\pi} (1 - \cos y) p(y)\, dy\Big]^{\frac{1}{2}} = \frac{\pi}{2} \sqrt{2 - c_1}.
\end{aligned}$$

Therefore,

$$\|\mathrm{K}f - f\|_\infty \leqslant \Big(1 + \frac{n\pi}{2} \sqrt{2 - c_1}\Big) \sup_{|x-y| \leqslant \frac{1}{n}} \big|f(x) - f(y)\big|.$$

To conclude the theorem, we need to show that the number $n\sqrt{2 - c_1}$ can be made bounded by choosing p properly. Nevertheless, let

$$\begin{aligned}
p(x) &= c \Big| \sum_{k=0}^{n} \sin \frac{(k+1)\pi}{n+2} e^{ikx} \Big|^2 = c \sum_{k=0}^{n} \sum_{\ell=0}^{n} \sin \frac{(k+1)\pi}{n+2} \sin \frac{(\ell+1)\pi}{n+2} e^{i(k-\ell)x} \\
&= c \sum_{k=0}^{n} \sin^2 \frac{(k+1)\pi}{n+2} + 2c \sum_{\substack{k,\ell=0 \\ k>\ell}}^{n} \sin \frac{(k+1)\pi}{n+2} \sin \frac{(\ell+1)\pi}{n+2} \cos(k-\ell)x
\end{aligned}$$

and choose c so that $p(x) = 1 + c_1 \cos x + \cdots + c_n \cos nx$. Then

$$\begin{aligned} c^{-1} &= \sum_{k=0}^{n} \sin^2 \frac{(k+1)\pi}{n+2} = \frac{1}{2} \sum_{k=0}^{n} \Big[1 - \cos \frac{2(k+1)\pi}{n+2}\Big] \\ &= \frac{n+1}{2} - \frac{\sin \frac{(2n+3)\pi}{n+2} - \sin \frac{\pi}{n+2}}{4 \sin \frac{\pi}{n+2}} = \frac{n+2}{2}, \end{aligned}$$

and

$$\begin{aligned} c_1 &= 2c \sum_{k=1}^{n} \sin \frac{(k+1)\pi}{n+2} \sin \frac{k\pi}{n+2} = c \sum_{k=1}^{n} \Big[\cos \frac{\pi}{n+2} - \cos \frac{(2k+1)\pi}{n+2}\Big] \\ &= c\Big[n \cos \frac{\pi}{n+2} - \frac{\sin \frac{(2n+2)\pi}{n+2} - \sin \frac{2\pi}{n+2}}{2 \sin \frac{\pi}{n+2}}\Big] \\ &= c\Big[n \cos \frac{\pi}{n+2} + \frac{\sin \frac{2\pi}{n+2}}{\sin \frac{\pi}{n+2}}\Big] = c(n+2) \cos \frac{\pi}{n+2} = 2 \cos \frac{\pi}{n+2}. \end{aligned}$$

As a consequence,

$$\begin{aligned} n\sqrt{2 - c_1} &= n\Big(2 - 2\cos \frac{\pi}{n+2}\Big)^{\frac{1}{2}} = 2n \sin \frac{\pi}{2(n+2)} \\ &= 2(n+2) \sin \frac{\pi}{2(n+2)} - 4 \sin \frac{\pi}{2(n+2)} \\ &= \pi \frac{2(n+2)}{\pi} \sin \frac{\pi}{2(n+2)} - 4 \sin \frac{\pi}{2(n+2)} \end{aligned}$$

which is bounded by π; thus

$$\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty \leqslant \|\mathrm{K}f - f\|_\infty \leqslant \big(1 + \frac{\pi^2}{2}\big) \sup_{|x-y| \leqslant \frac{1}{n}} \big|f(x) - f(y)\big|. \qquad \square$$

Finally, since $\lim\limits_{n \to \infty} n^{-\alpha} \log n = 0$ for all $\alpha \in (0, 1]$, we conclude the following

Theorem 8.17. *For any $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ with $\alpha \in (0, 1]$, the Fourier series of f converges uniformly to f on $\mathbb{R}$.*

Remark 8.18. The converse of Theorem 8.16 is the Bernstein theorem which states that if f is a 2π-periodic function with the property that there exist a constant C (independent of n) and $\alpha \in (0, 1)$ such that

$$\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty \leqslant C n^{-\alpha} \qquad \forall\, n \in \mathbb{N}, \tag{8.2.8}$$

then $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$. In other words, (8.2.8) is an equivalent condition to the Hölder continuity with exponent α of 2π-periodic continuous functions.

8.3 Cesàro Mean of Fourier Series

While Corollary 7.65 shows that the collection of trigonometric polynomials

$$\Big\{\frac{c_0}{2}+\sum_{k=1}^{n}(c_k\cos kx+s_k\sin kx)\,\Big|\,\{c_k\}_{k=0}^{n},\{s_k\}_{k=1}^{n}\subseteq\mathbb{R}\Big\}$$

is dense in $\mathscr{C}(\mathbb{T})$, Theorem 8.17 only implies the uniform convergence of the Fourier series of Hölder continuous functions. To approximate continuous functions uniformly, the coefficients of the trigonometric polynomials should depend on the order of approximation.

The motivation of the discussion below is due to the following observation. Let $\{a_k\}_{k=1}^{\infty}$ be a sequence. Define a new sequence $\{b_n\}_{n=1}^{\infty}$, called the ***Cesàro mean*** of the sequence $\{a_k\}_{k=1}^{\infty}$, by

$$b_n=\frac{a_1+\cdots+a_n}{n}=\frac{1}{n}\sum_{k=1}^{n}a_k\,.$$

If $\{a_k\}_{k=1}^{\infty}$ converges to a, then $\{b_n\}_{n=1}^{\infty}$ converges to a as well. Even though the convergence of a sequence cannot be guaranteed by the convergence of its Cesàro mean, it is worthwhile investigating the convergence behavior of the Cesàro mean.

Let $\sigma_n(f,\cdot)$ denote the Cesàro mean of the Fourier series of f given by

$$\sigma_n(f,\cdot)\equiv\frac{1}{n+1}\sum_{k=0}^{n}s_k(f,\cdot)=\frac{1}{n+1}\sum_{k=0}^{n}(D_k\star f)=\Big(\frac{1}{n+1}\sum_{k=0}^{n}D_k\Big)\star f\,.$$

We note that the coefficients of the Cesàro mean $\sigma_n(f,\cdot)$ depend on the order of approximation n since

$$\sigma_n(f,x)=\frac{c_0}{2}+\sum_{k=1}^{n}\Big(\underbrace{\frac{n+1-k}{n+1}c_k}_{\equiv c_k^{(n)}}\cos kx+\underbrace{\frac{n+1-k}{n+1}s_k}_{\equiv s_k^{(n)}}\sin kx\Big)\,.$$

Recall that $D_k(x)=\dfrac{\sin(k+\frac{1}{2})x}{2\pi\sin\frac{x}{2}}$. By the product-to-sum formula, we find that if $x\in(0,\pi)$,

$$\begin{aligned}\sum_{k=0}^{n}D_k(x)&=\frac{1}{2\pi\sin\frac{x}{2}}\sum_{k=0}^{n}\sin(k+\frac{1}{2})x=\frac{1}{4\pi\sin^2\frac{x}{2}}\sum_{k=0}^{n}2\sin\frac{x}{2}\sin(k+\frac{1}{2})x\\&=\frac{1}{4\pi\sin^2\frac{x}{2}}\sum_{k=0}^{n}\big(\cos kx-\cos(k+1)x\big)\\&=\frac{1}{4\pi\sin^2\frac{x}{2}}\big(1-\cos(n+1)x\big)=\frac{\sin^2\frac{n+1}{2}x}{2\pi\sin^2\frac{x}{2}}\,.\end{aligned}$$

This induces the following

Definition 8.19. The ***Fejér kernel*** is the Cesàro mean of the Dirichlet kernel given by

$$F_n(x) = \frac{1}{n+1}\sum_{k=0}^{n} D_k(x) = \frac{1}{2\pi(n+1)}\frac{\sin^2\frac{(n+1)x}{2}}{\sin^2\frac{x}{2}}.$$

We note that $\sigma_n(f,\cdot) = F_n \star f$, where $F_n \geqslant 0$ and has the property that $\displaystyle\int_{-\pi}^{\pi} F_n(x)\,dx = 1$ (since the integral of the Dirichlet kernel is 1). Moreover, for any $\delta > 0$,

$$\lim_{n\to\infty}\int_{\delta\leqslant|x|\leqslant\pi} F_n(x)\,dx = 0 \tag{8.3.1}$$

since $\big|F_n(x)\big| \leqslant \dfrac{1}{2\pi(n+1)\sin^2\frac{\delta}{2}}$ if $\delta \leqslant |x| \leqslant \pi$. Inequality (8.3.1) allows us to show that $\big\{\sigma_n(f,\cdot)\big\}_{n=1}^{\infty}$ converges uniformly to f.

Theorem 8.20. *For any $f \in \mathscr{C}(\mathbb{T})$, the Cesàro mean $\big\{\sigma_n(f,\cdot)\big\}_{n=1}^{\infty}$ of the Fourier series of f converges uniformly to f.*

Proof. Let $\varepsilon > 0$ be given. Since $f \in \mathscr{C}(\mathbb{T})$, f is uniformly continuous on $\mathbb{R}$; thus there exists $\delta > 0$ such that

$$\big|f(x) - f(y)\big| < \frac{\varepsilon}{2} \quad \text{whenever} \quad |x-y| < \delta.$$

Therefore, by the fact that $\displaystyle\int_{-\pi}^{\pi} F_n(x)dx = 1$ and $F_n \geqslant 0$,

$$\begin{aligned}
\big|\sigma_n(f,x) - f(x)\big| &= \Big|\int_{-\pi}^{\pi} F_n(y)f(x-y)\,dy - \int_{-\pi}^{\pi} F_n(y)f(x)\,dy\Big| \\
&\leqslant \int_{-\pi}^{\pi} F_n(y)\big|f(x-y) - f(x)\big|\,dy \\
&= \int_{|y|<\delta} F_n(y)\big|f(x-y) - f(x)\big|\,dy + \int_{\delta\leqslant|y|\leqslant\pi} F_n(y)\big|f(x-y) - f(x)\big|\,dy \\
&\leqslant \varepsilon\int_{|y|<\delta} F_n(y)\,dy + 2\|f\|_\infty\int_{\delta\leqslant|y|\leqslant\pi} F_n(y)\,dy \\
&\leqslant \frac{\varepsilon}{2} + 2\|f\|_\infty\int_{\delta\leqslant|y|\leqslant\pi} F_n(y)\,dy.
\end{aligned}$$

Using (8.3.1), there exists $N > 0$ such that

$$2\|f\|_\infty\int_{\delta\leqslant|y|\leqslant\pi} F_n(y)\,dy < \frac{\varepsilon}{2} \quad \textit{whenever} \quad n \geqslant N.$$

Therefore, $\big|\sigma_n(f,x) - f(x)\big| < \varepsilon$ whenever $n \geqslant N$ and $x \in \mathbb{R}$; thus we conclude that the Cesàro mean $\big\{\sigma_n(f,\cdot)\big\}_{n=1}^{\infty}$ converges uniformly to f. □

8.4 Convergence of Fourier Series for Functions with Jump Discontinuity

In previous sections we discussed the convergence of the Fourier series of continuous functions. However, since the Fourier series can be defined for bounded Riemann integrable functions, it is natural to ask what happen if the function under consideration is not continuous. We note that in this case we cannot apply Corollary 7.65 at all so no uniform convergence is expected.

In this section, we focus on the convergence behavior of Fourier series of functions with only jump discontinuities.

Definition 8.21. A function $f : [-\pi, \pi] \to \mathbb{R}$ is said to have jump discontinuity at $a \in (-\pi, \pi)$ if

1. $\lim\limits_{x\to a^+} f(x)$ and $\lim\limits_{x\to a^-} f(x)$ both exist.
2. $\lim\limits_{x\to a^+} f(x) \neq \lim\limits_{x\to a^-} f(x)$.

Now suppose that $f : [-\pi, \pi] \to \mathbb{R}$ is piecewise Hölder continuous with exponent $\alpha \in (0, 1]$; that is, there exists $\{a_1, \cdots, a_m\} \subseteq (-\pi, \pi)$ such that $f \in \mathscr{C}^{0,\alpha}((a_j, a_{j+1}); \mathbb{R})$ for all $j \in \{0, \cdots, m\}$, where $a_0 = -\pi$ and $a_{m+1} = \pi$, and $f \in \mathscr{C}^{0,\alpha}(I; \mathbb{R})$ if and only if

$$\sup_{x,y\in I, x\neq y} \frac{\big|f(x) - f(y)\big|}{|x - y|^\alpha} < \infty\,.$$

Then for all $a \in (-\pi, \pi)$, the limits $\lim\limits_{x\to a^+} f(x)$ and $\lim\limits_{x\to a^-} f(x)$ exist since if $\{x_k\}_{k=1}^\infty$ is a sequence in $(-\pi, \pi)$ which approaches to a from the right/left, then for some $0 \leqslant j \leqslant m$ we must have $x_k \in (a_j, a_{j+1})$ for all large k so that the Hölder continuity implies that

$$\big|f(x_k) - f(x_\ell)\big| \leqslant M|x_k - x_\ell|^\alpha \qquad \forall\, k, \ell \text{ large}$$

which shows that $\{f(x_k\}_{k=1}^\infty$ is a Cauchy sequence $\big($converging to $\lim\limits_{x\to a^\pm} f(x)\big)$. In other words, if $f : [-\pi, \pi] \to \mathbb{R}$ is piecewise Hölder continuous and $a \in (-\pi, \pi)$ is a discontinuity of f, then f has either removable discontinuity at a $\big($which means $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^-} f(x) \neq f(a)\big)$ or jump discontinuity at a. In the following, we always assume that f is piecewise Hölder continuous with exponent $\alpha \in (0, 1]$ and has only jump discontinuities at $\{a_1, \cdots, a_m\}$ in $(-\pi, \pi)$.

Let $f(a_j^+) = \lim\limits_{x \to a_j^+} f(x)$, $f(a_j^-) = \lim\limits_{x \to a_j^-} f(x)$, and define $\phi : \mathbb{R} \to \mathbb{R}$ by

$$\phi(x) = \frac{1}{2\pi}(x - \pi) \qquad \forall\, x \in [0, 2\pi) \tag{8.4.1}$$

and $\phi(x + 2\pi) = \phi(x)$ for all $x \in \mathbb{R}$. Since f has jump discontinuities at $\{a_1, \cdots, a_m\}$, with a_0^- denoting a_{m+1}^- the function $g : [-\pi, \pi] \to \mathbb{R}$ defined by

$$g(x) \equiv \begin{cases} f(x) + \sum\limits_{j=0}^{m} \big(f(a_j^+) - f(a_j^-)\big)\phi(x - a_j) & \text{if } x \neq a_k \text{ for all } k, \\ \dfrac{f(a_k^+) + f(a_k^-)}{2} + \sum\limits_{\substack{0 \leqslant j \leqslant m \\ j \neq k}} \big(f(a_j^+) - f(a_j^-)\big)\phi(a_k - a_j) & \text{if } x = a_k \text{ for some } k, \end{cases} \tag{8.4.2}$$

is Hölder continuous with exponent α and $g(a_0^+) = g(a_0^-) = g(-\pi)$. Let G be the 2π-periodic extension of g; that is, $G = g$ on $[-\pi, \pi]$ and $G(x + 2\pi) = G(x)$ for all $x \in \mathbb{R}$. Then $G \in \mathscr{C}^{0,\alpha}(\mathbb{T})$; thus Theorem 8.17 implies that $s_n(G, \cdot) \to G$ uniformly on $\mathbb{R}$. In particular, $s_n(g, \cdot) \to g$ uniformly on $[-\pi, \pi]$.

Using the identity

$$\int_{-\pi}^{\pi} \phi(x - a) e^{-ikx}\, dx = e^{-ika} \int_{-\pi}^{\pi} \phi(x) e^{-ikx}\, dx = \widehat{\phi}_k e^{-ika},$$

we obtain that

$$s_n(\phi(\cdot - a), x) = \sum_{k=-n}^{n} \widehat{\phi}_k e^{ik(x-a)} = s_n(\phi, x - a); \tag{8.4.3}$$

thus (8.4.2) implies that the Fourier series of f is given by

$$\begin{aligned} s_n(f, x) &= s_n(g, x) - \sum_{j=0}^{m} \big(f(a_j^+) - f(a_j^-)\big) s_n(\phi(\cdot - a_j), x) \\ &= s_n(g, x) - \sum_{j=0}^{m} \big(f(a_j^+) - f(a_j^-)\big) s_n(\phi, x - a_j). \end{aligned} \tag{8.4.4}$$

Therefore, to understand the convergence of the Fourier series of f, without loss of generality it suffices to consider the convergence of $s_n(\phi, \cdot)$.

8.4.1 Uniform convergence on compact subsets

In this sub-section, we show that the Fourier series of a piecewise Hölder continuous function whose discontinuities are all jump discontinuities converges uniformly on each compact subset containing no jump discontinuities.

Based on the discussion above, we first study the convergence of $s_n(\phi,\cdot)$. Since ϕ is an odd function, for $k \in \mathbb{N}$,

$$\begin{aligned} s_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} \phi(x)\sin kx\,dx = \frac{1}{\pi^2}\int_0^{\pi}(x-\pi)\sin kx\,dx \\ &= \frac{1}{\pi^2}\Big[\frac{-(x-\pi)\cos kx}{k}\Big|_{x=0}^{x=\pi} + \int_0^{\pi}\frac{\cos kx}{k}\,dx\Big] = -\frac{1}{\pi k}. \end{aligned}$$

Therefore, the n-th partial sum of the Fourier series of ϕ is given by

$$s_n(\phi,x) = -\frac{1}{\pi}\sum_{k=1}^{n}\frac{\sin kx}{k}. \tag{8.4.5}$$

Lemma 8.22. *The series* $\sum\limits_{k=1}^{\infty}\frac{\sin kx}{k}$ *converges uniformly on* $[-\pi,-\delta]\cup[\delta,\pi]$ *for all* $0<\delta<\pi$*.*

Proof. Let $0<\delta<\pi$ be given, and $S_n(x)$ denote the sum $\sum\limits_{k=1}^{n}\sin kx$. Using the identity

$$\sum_{k=1}^{n}\sin kx = \frac{\cos(n+\frac{1}{2})x - \cos\frac{x}{2}}{2\sin\frac{x}{2}} \qquad \forall\, x \in [-\pi,-\delta]\cup[\delta,\pi],$$

we find that $|S_n| \leqslant M < \infty$ for some fixed constant M. For $m>n$,

$$\begin{aligned} \sum_{k=n+1}^{m}\frac{1}{k}\sin kx &= \frac{1}{m}(S_m - S_{m-1}) + \frac{1}{m-1}(S_{m-1}-S_{m-2}) + \cdots + \frac{1}{n+1}(S_{n+1}-S_n) \\ &= \frac{S_m}{m} - \frac{S_n}{n+1} + \frac{1}{m(m-1)}S_{m-1} + \frac{1}{(m-1)(m-2)}S_{m-2} + \cdots \frac{1}{(n+2)(n+1)}S_{n+1}; \end{aligned}$$

thus

$$\Big|\sum_{k=n+1}^{m}\frac{1}{k}\sin kx\Big| \leqslant M\Big(\frac{1}{m}+\frac{1}{n+1}+\sum_{k=n+2}^{m}\frac{1}{k(k-1)}\Big) \leqslant 2M\Big(\frac{1}{m}+\frac{1}{n}\Big).$$

Since the right-hand side converges to 0 as $n,m\to\infty$, the Cauchy criterion (for the convergence of series of functions) implies that the series

$$\sum_{k=1}^{\infty}\frac{\sin kx}{k}$$

converges uniformly on $[-\pi,-\delta]\cup[\delta,\pi]$. ▫

Lemma 8.22 provides the uniform convergence of $s_n(\phi,\cdot)$ in $[-\pi,-\delta]\cup[\delta,\pi]$. To see the limit is exactly ϕ, we consider an anti-derivative Φ of ϕ and establish that $\Phi' = s(\phi,\cdot)$.

Let $\Psi : \mathbb{R} \to \mathbb{R}$ be 2π-periodic and $\Psi(x) = \dfrac{x^2}{4\pi}$ for $x \in [-\pi, \pi]$. Then $\Psi \in \mathscr{C}^{0,1}(\mathbb{T})$ is an even function and the Fourier coefficients of Ψ is

$$\widehat{\Psi}_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{x^2}{4\pi}\,dx = \frac{\pi}{12}$$

and for $k \neq 0$,

$$\widehat{\Psi}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{x^2}{4\pi} e^{-ikx}\,dx = \frac{1}{8\pi^2}\int_{-\pi}^{\pi} x^2(\cos kx + i\sin kx)\,dx = \frac{(-1)^k}{2k^2\pi}.$$

Therefore, using (8.4.3) we find that the Fourier series of $\Phi \equiv \Psi(\cdot - \pi)$ is

$$\begin{aligned} s(\Phi, x) = s(\Psi, x-\pi) &= \frac{\pi}{12} + \sum_{k\in\mathbb{Z}, k\neq 0} \widehat{\Psi}_k e^{ik(x-\pi)} = \frac{\pi}{12} + \frac{1}{2\pi}\sum_{k\in\mathbb{Z}, k\neq 0} \frac{e^{ikx}}{k^2} \\ &= \frac{\pi}{12} + \frac{1}{\pi}\sum_{k=1}^{\infty} \frac{\cos kx}{k^2}. \end{aligned}$$

Since $\Phi \in \mathscr{C}^{0,1}(\mathbb{T})$, $s_n(\Phi, \cdot)$ converges uniformly to Φ on $\mathbb{R}$. Moreover, $s_n(\Phi, \cdot)' = s_n(\phi, \cdot)$ which converges uniformly on $[-\pi, -\delta] \cup [\delta, \pi]$. Therefore, Theorem 7.11 implies that $s(\phi, \cdot)$, the uniform limit of $s_n(\phi, \cdot)$, must equal Φ' on $[-\pi, -\delta] \cup [\delta, \pi]$. Finally, we note that $\phi = \Phi'$ on $[-\pi, -\delta] \cup [\delta, \pi]$, so we establish that $s_n(\phi, \cdot) \to \phi$ uniformly on $[-\pi, -\delta] \cup [\delta, \pi]$.

Since a discontinuity of a piecewise Hölder continuous function f is either removable or a jump discontinuity, and the value of the function at removable discontinuities does not change the value of the Fourier series of f, the uniform convergence of $s_n(\phi, \cdot)$ to ϕ on $[-\pi, -\delta] \cup [\delta, \pi]$ for all $0 < \delta < \pi$ implies the following

Theorem 8.23. *Let $f : (-\pi, \pi) \to \mathbb{R}$ be piecewise Hölder continuous with exponent $\alpha \in (0, 1]$. If f is continuous on (a, b), then the Fourier series of f converges uniformly to f on any compact subsets of (a, b).*

By Remark 8.3, we can also conclude the following

Corollary 8.24. *Let $f : (-L, L) \to \mathbb{R}$ be piecewise Hölder continuous with exponent $\alpha \in (0, 1]$. If f is continuous on (a, b), then the Fourier series of f converges uniformly to f on any compact subsets of (a, b) (where the Fourier series of f is given in Remark 8.3). In particular, $\lim\limits_{n\to\infty} s_n(f, x_0) = f(x_0)$ if f is continuous at x_0. In other words, the Fourier series of f converges pointwise to f except the discontinuities.*

8.4.2 Gibbs phenomenon

In this sub-section, we show that the Fourier series evaluated at the jump discontinuity converges to the average of the limits from the left and the right. Moreover, the convergence of the Fourier series is never uniform in the domain including these jump discontinuities due to the famous Gibbs phenomenon: near the jump discontinuity the maximum difference between the limit of the Fourier series and the function itself is at least 8% of the jump. To be more precise, we have the following

Theorem 8.25. *Let $f:\mathbb{R}\to\mathbb{R}$ be $2L$-periodic piecewise Hölder continuous with exponent $\alpha\in(0,1]$. Then*

$$\lim_{n\to\infty} s_n(f,x_0)=\frac{f(x_0^+)+f(x_0^-)}{2}\qquad \forall\, x_0\in\mathbb{R}. \tag{8.4.6}$$

Moreover, if x_0 is a jump discontinuity of f so that

$$f(x_0^+)-f(x_0^-)=a\neq 0\,,$$

then there exists a constant $c>0$, independent of f, x_0 and L (in fact, $c=\dfrac{1}{\pi}\displaystyle\int_0^\pi \frac{\sin x}{x}\,dx-\frac{1}{2}\approx 0.089490$), such that

$$\lim_{n\to\infty} s_n\big(f,x_0+\frac{L}{n}\big)=f(x_0^+)+ca\,, \tag{8.4.7a}$$

$$\lim_{n\to\infty} s_n\big(f,x_0-\frac{L}{n}\big)=f(x_0^-)-ca\,. \tag{8.4.7b}$$

Proof. By Remark 8.3, W.L.O.G. we can assume that $L=\pi$. Let $\{a_1,\cdots,a_m\}\subseteq(-\pi,\pi)$ be the collection of jump discontinuities of f in $(-\pi,\pi)$, $a_0=-\pi$, $a_{m+1}=\pi$ (so by periodicity $f(a_0^-)=f(a_{m+1}^-)$ automatically), and define g by (8.4.2). Then $g\in\mathscr{C}^{0,\alpha}(\mathbb{T})$. Suppose that x_0 is a jump discontinuity of f in $[-\pi,\pi)$ (so a_0 could be a possible jump discontinuity of f). Then $x_0=a_k$ for some $k\in\{0,1,\cdots,m\}$. Therefore, by the fact that ϕ is continuous at x_0-a_j if $j\neq k$ and $s_n(\phi,0)=0$ for all $n\in\mathbb{N}$, Corollary 8.24 implies that

$$\begin{aligned}&\sum_{j=0}^m \big(f(a_j^+)-f(a_j^-)\big)\lim_{n\to\infty} s_n(\phi,x_0-a_j)\\ &\qquad=\sum_{\substack{0\leqslant j\leqslant m\\ j\neq k}}\big(f(a_j^+)-f(a_j^-)\big)\lim_{n\to\infty} s_n(\phi,x_0-a_j)=\sum_{\substack{0\leqslant j\leqslant m\\ j\neq k}}\big(f(a_j^+)-f(a_j^-)\big)\phi(x_0-a_j)\,.\end{aligned}$$

On the other hand,

$$\lim_{n\to\infty} s_n(g,x_0)=g(x_0)=\frac{f(x_0^+)+f(x_0^-)}{2}+\sum_{\substack{0\leqslant j\leqslant m\\ j\neq k}}\big(f(a_j^+)-f(a_j^-)\big)\phi(x_0-a_j)\,.$$

Identity (8.4.6) is then concluded using (8.4.4).

Now we focus on (8.4.7a). Since $g \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, $s_n(g,\cdot) \to g$ uniformly on $\mathbb{R}$; thus

$$\lim_{n\to\infty} s_n\big(g, x_0 + \frac{\pi}{n}\big) = g(x_0)\,.$$

Similarly, since $s_n(\phi,\cdot) \to \phi$ uniformly on $[-\pi,-\delta] \cup [\delta,\pi]$ for all $\delta > 0$, if $j \neq k$,

$$\lim_{n\to\infty} s_n\big(\phi, x_0 + \frac{\pi}{n} - a_j\big) = \phi(x_0 - a_j)\,.$$

On the other hand,

$$s_n\big(\phi, \frac{\pi}{n}\big) = -\sum_{k=1}^{n} \frac{1}{\pi k} \sin\frac{k\pi}{n} = -\frac{1}{\pi}\sum_{k=1}^{n} \frac{n}{k\pi}\sin\frac{k\pi}{n}\frac{\pi}{n} \to -\frac{1}{\pi}\int_0^{\pi} \frac{\sin x}{x} dx \equiv -\big(c + \frac{1}{2}\big)\,.$$

As a consequence,

$$\begin{aligned}
\lim_{n\to\infty} s_n\big(f, x_0 + \frac{\pi}{n}\big) &= \lim_{n\to\infty}\Big[s_n\big(g, x_0 + \frac{\pi}{n}\big) - \sum_{j=0}^{m}\big(f(a_j^+) - f(a_j^-)\big)s_n\big(\phi, x_0 + \frac{\pi}{n} - a_j\big)\Big] \\
&= g(x_0) - \sum_{\substack{0\leqslant j\leqslant m \\ j\neq k}}\big(f(a_j^+) - f(a_j^-)\big)\phi(x_0 - a_j) + \big(c + \frac{1}{2}\big)\big(f(x_0^+) - f(x_0^-)\big) \\
&= f(x_0^+) + c\big(f(x_0^+) - f(x_0^-)\big)\,.
\end{aligned}$$

Identity (8.4.7b) can be proved in the same fashion, and is left as an exercise. □

Remark 8.26. Let f be a function given in Theorem 8.25, x_0 be a jump discontinuity of f, and $I = (x_0, x_0 + r)$ for some $r > 0$ so that f is continuous on I. By the definition of the right limit, there exists $0 < \delta < r$ such that

$$\big|f(x) - f(x_0^+)\big| < \frac{c|a|}{2} \qquad \forall\, x \in (x_0, x_0 + \delta)\,.$$

Choose $N > 0$ such that $\dfrac{L}{N} < \delta$. Then $x_0 + \dfrac{L}{N} \in (x_0, x_0 + \delta)$ for all $n \geqslant N$; thus if $n \geqslant N$,

$$\begin{aligned}
\sup_{x\in I}\big|s_n(f,x) - f(x)\big| &\geqslant \big|s_n(f, x_0 + \frac{L}{N}) - f(x_0 + \frac{L}{N})\big| \\
&\geqslant \big|s_n(f, x_0 + \frac{L}{N}) - f(x_0^+)\big| - \big|f(x_0 + \frac{L}{N}) - f(x_0^+)\big| \\
&\geqslant \big|s_n(f, x_0 + \frac{L}{N}) - f(x_0^+)\big| - \frac{c|a|}{2}
\end{aligned}$$

which implies that

$$\liminf_{n\to\infty}\sup_{x\in I}\big|s_n(f,x)-f(x)\big| \geqslant c|a| - \frac{c|a|}{2} = \frac{c|a|}{2}\,.$$

Therefore, $\big\{s_n(f,\cdot)\big\}_{n=1}^{\infty}$ does not converge uniformly (to f) on I, while Corollary 8.24 shows that $\big\{s_n(f,\cdot)\big\}_{n=1}^{\infty}$ converges pointwise to f on I. Similarly, if x_0 is a jump discontinuity of f and f is continuous on (x_0-r,x_0) for some $r>0$, then $\big\{s_n(f,\cdot)\big\}_{n=1}^{\infty}$ converge pointwise but not uniformly on (x_0-r,x_0).

For a function f given in Theorem 8.25, let $\widetilde{f}$ be defined by

$$\widetilde{f}(x) = \left\{\begin{array}{cl} f(x) & \text{if } f \text{ is continuous at } x\,, \\ \dfrac{f(x^+)+f(x^-)}{2} & \text{if } x \text{ is a discontinuity of } f\,. \end{array}\right.$$

Then $s_n(\widetilde{f},\cdot) = s_n(f,\cdot)$ for all $n\in\mathbb{N}$, and Corollary 8.24 and Theorem 8.25 together imply that $\big\{s_n(f,\cdot)\big\}_{n=1}^{\infty}$ converges pointwise to $\widetilde{f}$. However, the discussion above shows that $\big\{s_n(f,\cdot)\big\}_{n=1}^{\infty}$ cannot converge uniformly on I as long as I contains jump discontinuities of f.

8.5 The Inner-Product Point of View

除了逐點收斂或均勻收斂的觀點之外，還有一個更自然（就數學而言）的觀點可以用來看 Fourier series。我們可以把定義在 $[-\pi,\pi]$ 的所有 square integrable 函數（定義在下）所形成的集合看成一個向量空間，然後在上面定義一個內積的結構。一個可積分函數（也可視為一個向量）的 Fourier series 可以看成這個向量在一組正交基底向量的線性組合。

Let $L^2(\mathbb{T})$ denote the collection of Riemann measurable, square integrable function on $[-\pi,\pi]$ modulo the relation that $f\sim g$ if $f-g=0$ except on a set of measure zero (or $f=g$ almost everywhere). In other words,

$$L^2(\mathbb{T}) = \Big\{f:[-\pi,\pi]\to\mathbb{C}\,\Big|\, f \text{ is Riemann measurable and } \int_{-\pi}^{\pi}|f(x)|^2\,dx<\infty\Big\}\Big/\sim\,.$$

Here we **abuse** the use of notation $L^2(\mathbb{T})$ for that it indeed denotes a more general space. We also note that the domain $[-\pi,\pi]$ can be replaced by any intervals with $-\pi$, π as end-points for we can easily modify functions defined on those domains to functions defined on $[-\pi,\pi]$ without changing the square integrability.

Define a bilinear function $\langle\cdot,\cdot\rangle$ on $L^2(\mathbb{T})\times L^2(\mathbb{T})$ by

$$\langle f,g\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)\overline{g(x)}\,dx\,.$$

Then $\langle\cdot,\cdot\rangle$ is an inner product on $L^2(\mathbb{T})$. Indeed, if f,g belong to $L^2(\mathbb{T})$, then the product $f\bar{g}$ is also Riemann measurable, and the Cauchy-Schwartz inequality as well as the monotone convergence theorem imply that

$$\begin{aligned}\big|\langle f,g\rangle\big|&=\lim_{k\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}\big|(f\wedge k)(x)\big|\big|(g\wedge k)(x)\big|\,dx\\&\leqslant\lim_{k\to\infty}\frac{1}{2\pi}\Big(\int_{-\pi}^{\pi}\big|(f\wedge k)(x)\big|^2\,dx\Big)^{\frac12}\Big(\int_{-\pi}^{\pi}\big|(g\wedge k)(x)\big|^2\,dx\Big)^{\frac12}\\&=\Big(\frac{1}{2\pi}\int_{-\pi}^{\pi}\big|f(x)\big|^2\,dx\Big)^{\frac12}\Big(\frac{1}{2\pi}\int_{-\pi}^{\pi}\big|g(x)\big|^2\,dx\Big)^{\frac12}=\|f\|_{L^2(\mathbb{T})}\|g\|_{L^2(\mathbb{T})}<\infty;\end{aligned}$$

thus the definition of the inner product $\langle\cdot,\cdot\rangle$ given above is well-defined. The norm induced by the inner product above is denoted by $\|\cdot\|_{L^2(\mathbb{T})}$.

For $k\in\mathbb{Z}$, define $\mathbf{e}_k:[-\pi,\pi]\to\mathbb{C}$ by $\mathbf{e}_k(x)=e^{ikx}$. Then $\{\mathbf{e}_k\}_{k=-\infty}^{\infty}$ is an orthonormal set in $L^2(\mathbb{T})$ since

$$\langle\mathbf{e}_k,\mathbf{e}_\ell\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{ikx}e^{-i\ell x}\,dx=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i(k-\ell)x}\,dx=\begin{cases}1&\text{if }k=\ell,\\0&\text{if }k\neq\ell.\end{cases}$$

Let $\mathcal{V}_n=\operatorname{span}(\mathbf{e}_{-n},\mathbf{e}_{-n+1},\cdots,\mathbf{e}_0,\mathbf{e}_1,\cdots,\mathbf{e}_n)=\Big\{\sum\limits_{k=-n}^{n}a_k\mathbf{e}_k\,\Big|\,\{a_k\}_{k=-n}^{n}\subseteq\mathbb{C}\Big\}$. For each vector $f\in L^2(\mathbb{T})$, the orthogonal projection of f onto $\mathcal{V}_n$ is, conceptually, given by

$$\sum_{k=-n}^{n}\langle f,\mathbf{e}_k\rangle\mathbf{e}_k=\sum_{k=-n}^{n}\Big(\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)e^{-ikx}\,dx\Big)\mathbf{e}_k=\sum_{k=-n}^{n}\widehat{f}_k\mathbf{e}_k\,.$$

By the definition of $\mathbf{e}_k$, we obtain that the orthogonal projection of f on $\mathcal{V}_n$ is exactly $s_n(f,\cdot)$. We also note that $\mathcal{V}_n=\mathscr{P}_n(\mathbb{T})$.

Now we prove that $s_n(f,\cdot)$ is exactly the orthogonal projection of f onto $\mathcal{V}_n=\mathscr{P}_n(\mathbb{T})$.

Proposition 8.27. *Let $f\in L^2(\mathbb{T})$. Then*

$$\langle f-s_n(f,\cdot),p\rangle=0\qquad\forall\,p\in\mathscr{P}_n(\mathbb{T})\,.$$

Proof. Let $p\in\mathscr{P}_n(\mathbb{T})$. Then $p=s_n(p,\cdot)$; thus

$$\begin{aligned}\langle f-s_n(f,\cdot),p\rangle&=\langle f,p\rangle-\langle s_n(f,\cdot),p\rangle=\Big\langle f,\sum_{k=-n}^{n}\widehat{p}_k\mathbf{e}_k\Big\rangle-\Big\langle\sum_{k=-n}^{n}\widehat{f}_k\mathbf{e}_k,p\Big\rangle\\&=\sum_{k=-n}^{n}\overline{\widehat{p}_k}\langle f,\mathbf{e}_k\rangle-\sum_{k=-n}^{n}\widehat{f}_k\overline{\langle p,\mathbf{e}_k\rangle}=\sum_{k=-n}^{n}\overline{\widehat{p}_k}\widehat{f}_k-\sum_{k=-n}^{n}\widehat{f}_k\overline{\widehat{p}_k}=0\,.\end{aligned}$$

□

Theorem 8.28. *Let* $f \in L^2(\mathbb{T})$. *Then*

$$\|f - p\|^2_{L^2(\mathbb{T})} = \|f - s_n(f,\cdot)\|^2_{L^2(\mathbb{T})} + \|s_n(f,\cdot) - p\|^2_{L^2(\mathbb{T})} \qquad \forall\, p \in \mathscr{P}_n(\mathbb{T}). \tag{8.5.1}$$

Proof. By Proposition 8.27, if $p \in \mathscr{P}_n(\mathbb{T})$, $s_n(f,\cdot) - p = s_n(f-p,\cdot) \in \mathscr{P}_n(\mathbb{T})$; thus

$$\begin{aligned}
\|f - p\|^2_{L^2(\mathbb{T})} &= \langle f - p, f - p\rangle = \langle f - s_n(f,\cdot) + s_n(f,\cdot) - p, f - s_n(f,\cdot) + s_n(f,\cdot) - p\rangle \\
&= \big\|f - s_n(f,\cdot)\big\|^2_{L^2(\mathbb{T})} + 2\mathrm{Re}\big(\langle f - s_n(f,\cdot), s_n(f,\cdot) - p\rangle\big) + \big\|s_n(f,\cdot) - p\big\|^2_{L^2(\mathbb{T})} \\
&= \big\|f - s_n(f,\cdot)\big\|^2_{L^2(\mathbb{T})} + \big\|s_n(f,\cdot) - p\big\|^2_{L^2(\mathbb{T})}
\end{aligned}$$

which concludes the proposition. □

We note that (8.5.1) implies that

$$\|f - s_n(f,\cdot)\|_{L^2(\mathbb{T})} \leqslant \|f - p\|_{L^2(\mathbb{T})} \qquad \forall\, p \in \mathscr{P}_n(\mathbb{T}). \tag{8.5.2}$$

Since $s_n(f,\cdot) \in \mathscr{P}_n(\mathbb{T})$, we conclude that

$$\|f - s_n(f,\cdot)\|_{L^2(\mathbb{T})} = \inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f - p\|_{L^2(\mathbb{T})}.$$

Moreover, letting $p = 0$ in (8.5.1) we establish the famous Bessel's inequality.

Corollary 8.29. *Let* $f \in L^2(\mathbb{T})$. *Then for all* $n \in \mathbb{N}$,

$$\|s_n(f,\cdot)\|_{L^2(\mathbb{T})} \leqslant \|f\|_{L^2(\mathbb{T})}. \tag{8.5.3}$$

In particular,

$$\sum_{k=-\infty}^{\infty} |\widehat{f}_k|^2 \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} \big|f(x)\big|^2\,dx. \tag{Bessel's inequality}$$

Remark 8.30. When $f \in L^2(\mathbb{T})$ and f is real-valued, then

$$\sum_{k=-\infty}^{\infty} |\widehat{f}_k|^2 = \frac{c_0^2}{4} + \frac{1}{2}\sum_{k=1}^{\infty}(c_k^2 + s_k^2);$$

thus in this case the Bessel inequality can also be written as

$$\frac{c_0^2}{4} + \frac{1}{2}\sum_{k=1}^{\infty}(c_k^2 + s_k^2) \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} \big|f(x)\big|^2\,dx.$$

Next, we prove that the Bessel inequality is in fact an equality, called the Parseval identity. Using (8.5.1), it is equivalent to that $\big\{s_n(f,\cdot)\big\}_{n=1}^\infty$ converges to f in the sense of L^2-norm; that is,

$$\lim_{n\to\infty} \big\|s_n(f,\cdot)-f\big\|_{L^2(\mathbb{T})}=0 \qquad \forall\, f\in L^2(\mathbb{T})\,.$$

Before proceeding, we first prove that every $f \in L^2(\mathbb{T})$ can be approximated by a sequence $\{g_n\}_{n=1}^\infty \subseteq \mathscr{C}(\mathbb{T})$ in the sense of L^2-norm.

Proposition 8.31. *Let $f\in L^2(\mathbb{T})$. Then for all $\varepsilon>0$ there exists $g\in\mathscr{C}(\mathbb{T})$ (here $\mathscr{C}(\mathbb{T})$ denotes the collections of 2π-periodic complex-valued continuous functions on $\mathbb{R}$) such that*

$$\|f-g\|_{L^2(\mathbb{T})}<\varepsilon\,.$$

In other words, $\mathscr{C}(\mathbb{T})$ is dense in $\big(L^2(\mathbb{T}),\|\cdot\|_{L^2(\mathbb{T})}\big)$.

Proof. W.L.O.G., we can assume that f is real-valued and non-zero. Let $\varepsilon>0$ be given. Since $f\in L^2(\mathbb{T})$, the monotone convergence theorem (Corollary 6.105) implies that

$$\lim_{k\to\infty}\|f-(-k)\vee(f\wedge k)\|^2_{L^2(\mathbb{T})}=\lim_{k\to\infty}\int_{-\pi}^{\pi}\mathbf{1}_{\{|f(x)|>k\}}(x)|f(x)|^2\,dx=0;$$

thus there exists $N>0$ such that

$$\|f-(-k)\vee(f\wedge k)\|_{L^2(\mathbb{T})}<\frac{\varepsilon}{2}\qquad \forall\, k\geqslant N\,.$$

Let $h=(-N)\vee(f\wedge N)$. Then h is bounded and Riemann measurable; thus h is Riemann integrable on $[-\pi,\pi]$. Therefore, there exists a partition $\mathcal{P}=\{-\pi=x_0<x_1<\cdots<x_n=\pi\}$ of $[-\pi,\pi]$ such that $U(h,\mathcal{P})-L(h,\mathcal{P})<\dfrac{\pi\varepsilon^2}{8N}$. Define

$$S(x)=\sum_{k=0}^{n-1}\sup_{\xi\in[x_k,x_{k+1}]}h(\xi)\mathbf{1}_{[x_k,x_{k+1}]}(x)\quad\text{ and }\quad s(x)=\sum_{k=0}^{n-1}\inf_{\xi\in[x_k,x_{k+1}]}h(\xi)\mathbf{1}_{[x_k,x_{k+1}]}(x),$$

where $\mathbf{1}_A$ denotes the characteristic/indicator function of set A. Then

1. $-N\leqslant s\leqslant h\leqslant S\leqslant N$ on $[-\pi,\pi]\backslash\{x_1,x_2,\cdots,x_{n-1}\}$;
2. $\displaystyle\int_{-\pi}^{\pi}S(x)\,dx=U(h,\mathcal{P})$;
3. $\displaystyle\int_{-\pi}^{\pi}s(x)\,dx=L(h,\mathcal{P})$.

The properties above show that

$$\begin{aligned}\int_{-\pi}^{\pi}\big|h(x)-s(x)\big|\,dx&=\int_{-\pi}^{\pi}h(x)-s(x)\,dx\leqslant\int_{-\pi}^{\pi}\big(S(x)-s(x)\big)\,dx\\&=U(h,\mathcal{P})-L(h,\mathcal{P})<\frac{\pi\varepsilon^2}{8N}\,.\end{aligned}$$

Now, similar to the construction of g and h in the proof of Lemma 6.63, for the step function s defined on $[-\pi, \pi]$ we can always find a continuous function $g \in \mathscr{C}(\mathbb{T})$ such that

1. $\|g\|_{L^\infty(\mathbb{T})} \leqslant N$. 2. $\displaystyle\int_{-\pi}^{\pi} \big|s(x) - g(x)\big|\, dx < \frac{\pi\varepsilon^2}{8N}$.

Therefore,

$$\int_{-\pi}^{\pi} \big|h(x) - g(x)\big|\, dx \leqslant \int_{-\pi}^{\pi} \big|h(x) - s(x)\big|\, dx + \int_{-\pi}^{\pi} \big|s(x) - g(x)\big|\, dx < \frac{\pi\varepsilon^2}{4N}$$

which implies that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \big|h(x) - g(x)\big|^2\, dx \leqslant \frac{N}{\pi}\int_{[-\pi,\pi]} \big|h(x) - g(x)\big|\, dx < \frac{\varepsilon^2}{4};$$

thus $\|h - g\|_{L^2(\mathbb{T})} < \dfrac{\varepsilon}{2}$. The proposition is then concluded by the triangle inequality. □

Theorem 8.32. *Let $f \in L^2(\mathbb{T})$. Then*

$$\lim_{n\to\infty} \big\|f - s_n(f,\cdot)\big\|_{L^2(\mathbb{T})} = 0 \tag{8.5.4}$$

and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \big|f(x)\big|^2\, dx = \sum_{k=-\infty}^{\infty} |\widehat{f}_k|^2\,. \qquad \text{(Parseval's identity)}$$

Proof. Let $\varepsilon > 0$ be given. By Proposition 8.31 there exists $g \in \mathscr{C}(\mathbb{T})$ such that

$$\|f - g\|_{L^2(\mathbb{T})} < \frac{\varepsilon}{3}\,.$$

By the denseness of the trigonometric polynomials in $\mathscr{C}(\mathbb{T})$, there exists $h \in \mathscr{P}(\mathbb{T})$ such that $\sup\limits_{x\in\mathbb{R}} \big|g(x) - h(x)\big| < \dfrac{\varepsilon}{3}$. Suppose that $h \in \mathscr{P}_N(\mathbb{T})$. Using (8.5.2),

$$\big\|g - s_{\scriptscriptstyle N}(g,\cdot)\big\|^2_{L^2(\mathbb{T})} \leqslant \|g - h\|^2_{L^2(\mathbb{T})} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \big|g(x) - h(x)\big|^2\, dx \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\varepsilon^2}{9}\, dx = \frac{\varepsilon^2}{9}\,.$$

Since $s_{\scriptscriptstyle N}(g,\cdot) \in \mathscr{P}_n(\mathbb{T})$ if $n \geqslant N$, using (8.5.2) again we must have

$$\big\|g - s_n(g,\cdot)\big\|_{L^2(\mathbb{T})} \leqslant \big\|g - s_{\scriptscriptstyle N}(g,\cdot)\big\|_{L^2(\mathbb{T})} \leqslant \frac{\varepsilon}{3} \qquad \forall\, n \geqslant N\,.$$

Therefore, for $n \geqslant N$, inequality (8.5.3) and the triangle inequality yield that

$$\begin{aligned}\big\|f - s_n(f,\cdot)\big\|_{L^2(\mathbb{T})} &\leqslant \|f - g\|_{L^2(\mathbb{T})} + \|g - s_n(g,\cdot)\|_{L^2(\mathbb{T})} + \|s_n(g - f,\cdot)\|_{L^2(\mathbb{T})} \\ &\leqslant 2\|f - g\|_{L^2(\mathbb{T})} + \|g - s_n(g,\cdot)\|_{L^2(\mathbb{T})} < \varepsilon;\end{aligned}$$

thus (8.5.4) is concluded. Finally, using (8.5.1) with $p = 0$ we obtain that

$$\int_{-\pi}^{\pi} \big|f(x)\big|^2 \, dx = \int_{-\pi}^{\pi} \big|s_n(f,x)\big|^2 \, dx + \int_{-\pi}^{\pi} \big|f(x) - s_n(f,x)\big|^2 \, dx \, .$$

Using the fact that $\dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} \big|s_n(f,x)\big|^2 \, dx = \sum_{k=-n}^{n} \big|\widehat{f}_k\big|^2$ and passing to the limit as $n \to \infty$, we conclude the Parseval identity. □

Example 8.33. Example 8.6 provides that $\displaystyle\int_{-\pi}^{\pi} x^2 \, dx = \pi \sum_{k=1}^{\infty} \frac{4}{k^2}$; thus $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Remark 8.34. The Parseval identity implies that

$$\langle f, g \rangle_{L^2(\mathbb{T})} = \sum_{k=-\infty}^{\infty} \widehat{f}_k \overline{\widehat{g}_k} \qquad \forall \, f, g \in L^2(\mathbb{T}) \tag{8.5.5}$$

since the polarization identity shows that

$$\begin{aligned}
\langle f, g \rangle_{L^2(\mathbb{T})} &= \frac{1}{4}\Big[\|f+g\|^2_{L^2(\mathbb{T})} - \|f-g\|^2_{L^2(\mathbb{T})} + i\|f+ig\|^2_{L^2(\mathbb{T})} - i\|f-ig\|^2_{L^2(\mathbb{T})}\Big] \\
&= \frac{1}{4} \sum_{k=-\infty}^{\infty} \Big[|\widehat{f}_k + \widehat{g}_k|^2 - |\widehat{f}_k - \widehat{g}_k|^2 + i|\widehat{f}_k + i\widehat{g}_k|^2 - i|\widehat{f}_k - i\widehat{g}_k|^2\Big] \\
&= \frac{1}{4} \sum_{k=-\infty}^{\infty} \Big[\big(|\widehat{f}_k|^2 + 2\mathrm{Re}(\widehat{f}_k\overline{\widehat{g}_k}) + |\widehat{g}_k|^2\big) - \big(|\widehat{f}_k|^2 - 2\mathrm{Re}(\widehat{f}_k\overline{\widehat{g}_k} + |\widehat{g}_k|^2\big) \\
&\qquad\qquad + i\big(|\widehat{f}_k|^2 + 2\mathrm{Im}(\widehat{f}_k\overline{\widehat{g}_k}) + |\widehat{g}_k|^2\big) - i\big(|\widehat{f}_k|^2 - 2\mathrm{Im}(\widehat{f}_k\overline{\widehat{g}_k}) + |\widehat{g}_k|^2\big)\Big] \\
&= \sum_{k=-\infty}^{\infty} \big[\mathrm{Re}(\widehat{f}_k\overline{\widehat{g}_k}) + i\mathrm{Im}(\widehat{f}_k\overline{\widehat{g}_k})\big] = \sum_{k=-\infty}^{\infty} \widehat{f}_k\overline{\widehat{g}_k} \, .
\end{aligned}$$

8.6 The Discrete Fourier "Transform" and the Fast Fourier "Transform"

Let $f : \mathbb{R} \to \mathbb{R}$ be a periodic function with period L and f is bounded Riemann integrable on $[0, L)$. Similar to Remark 8.2, the Fourier series of f, defined in Remark 8.3, can be written as

$$s(f,x) = \sum_{k=-\infty}^{\infty} \widehat{f}_k e^{\frac{2\pi i k x}{L}} \, ,$$

where $\widehat{f}_k = \dfrac{1}{L}\displaystyle\int_0^L f(y)e^{\frac{-2\pi iky}{L}}\,dy$; thus $\widehat{f}_k$ can be approximated by the Riemann sum

$$\frac{1}{L}\sum_{\ell=0}^{N-1} f\big(\frac{L\ell}{N}\big)e^{\frac{-2\pi ik\ell}{N}}\frac{L}{N} = \frac{1}{N}\sum_{\ell=0}^{N-1} f\big(\frac{L\ell}{N}\big)e^{\frac{-2\pi ik\ell}{N}}.$$

In other words, the values of f at N evenly distributed points can be used to determine an approximation of the Fourier coefficients of f.

There is another point of view of finding the sum $\dfrac{1}{N}\displaystyle\sum_{\ell=0}^{N-1} f\big(\frac{L\ell}{N}\big)e^{\frac{-2\pi ik\ell}{N}}$. Even though $s_n(f,x)$ will be a good approximation of $s(f,x)$ for large n, the computation of the exact Fourier coefficients will be expensive (and probably impossible). Therefore, instead of compute the exact Fourier coefficients, we look for a Fourier-like series of the form

$$\frac{1}{N}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi ikx}{L}}.$$

so that it agrees with the value of f at points $\Big\{\dfrac{Lj}{N}\Big\}_{j=0}^{N-1}$. Therefore, we look for $\{X_k\}_{k=0}^{N-1}$ satisfying that

$$\frac{1}{N}\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & e^{\frac{2\pi i}{N}} & e^{\frac{4\pi i}{N}} & \cdots & e^{\frac{2\pi(N-1)i}{N}} \\ 1 & e^{\frac{4\pi i}{N}} & e^{\frac{8\pi i}{N}} & \cdots & e^{\frac{4\pi(N-1)i}{N}} \\ \vdots & & & \ddots & \vdots \\ 1 & e^{\frac{2\pi(N-1)i}{N}} & e^{\frac{4\pi(N-1)i}{N}} & \cdots & e^{\frac{2\pi(N-1)^2 i}{N}} \end{bmatrix}\begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ \vdots \\ X_{N-1} \end{bmatrix} = \begin{bmatrix} f(0) \\ f\big(\frac{L}{N}\big) \\ \vdots \\ f\big(\frac{(N-1)L}{N}\big) \end{bmatrix}.$$

Let $\boldsymbol{v}_k = \big[v_k^{(1)}, v_k^{(2)}, \cdots, v_k^{(N)}\big]^{\rm T}$ denote the k-th column of the $N\times N$ matrix F on the left-hand side of the equation above. Then $v_k^{(j)} = e^{\frac{2\pi(k-1)(j-1)i}{N}}$ so that

$$\begin{aligned}\boldsymbol{v}_\ell\cdot\boldsymbol{v}_k = \boldsymbol{v}_k^*\boldsymbol{v}_\ell &= \sum_{j=1}^{N} e^{\frac{-2\pi(j-1)(k-1)i}{N}} e^{\frac{2\pi(j-1)(\ell-1)i}{N}} = \sum_{j=1}^{N} e^{\frac{2\pi(j-1)(\ell-k)i}{N}} \\ &= \sum_{j=0}^{N-1}\cos\frac{2\pi j(\ell-k)}{N} + i\sum_{j=0}^{N-1}\sin\frac{2\pi j(\ell-k)i}{N}\end{aligned}$$

which shows that

$$\boldsymbol{v}_\ell\cdot\boldsymbol{v}_k = \begin{cases} N & \text{if } k=\ell, \\ 0 & \text{if } k\neq\ell. \end{cases}$$

Therefore, $F^*F = N\mathrm{I}_{N\times N}$; thus

$$\begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ \vdots \\ X_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & e^{-\frac{2\pi i}{N}} & e^{-\frac{4\pi i}{N}} & \cdots & e^{-\frac{2\pi(N-1)i}{N}} \\ 1 & e^{-\frac{4\pi i}{N}} & e^{-\frac{8\pi i}{N}} & \cdots & e^{-\frac{4\pi(N-1)i}{N}} \\ \vdots & & & \ddots & \vdots \\ 1 & e^{-\frac{2\pi(N-1)i}{N}} & e^{-\frac{4\pi(N-1)i}{N}} & \cdots & e^{-\frac{2\pi(N-1)^2 i}{N}} \end{bmatrix} \begin{bmatrix} f(0) \\ f\big(\frac{L}{N}\big) \\ \vdots \\ f\big(\frac{(N-1)L}{N}\big) \end{bmatrix}.$$

The discussions above induce the following

Definition 8.35. The ***discrete Fourier transform***, symbolized by DFT, of a sequence of N complex numbers $\{x_0, x_1, \cdots, x_{N-1}\}$ is a sequence $\{X_k\}_{k\in\mathbb{Z}}$ defined by

$$X_k = \sum_{\ell=0}^{N-1} x_\ell e^{\frac{-2\pi i k\ell}{N}} \qquad \forall\, k \in \mathbb{Z}.$$

We note that the sequence $\{X_k\}_{k\in\mathbb{Z}}$ is N-periodic; that is, $X_{k+N} = X_k$ for all $k \in \mathbb{Z}$. Therefore, often time we only focus on one of the following N consecutive terms $\{X_0, X_1, \cdots, X_{N-1}\}$ of the DFT.

Example 8.36. The DFT of the sequence $\{x_0, x_1\}$ is $\{x_0 + x_1, x_0 - x_1\}$.

8.6.1 The inversion formula

Let $\{X_k\}_{k=0}^{N-1}$ be the discrete Fourier transform of the sequence $\{x_\ell\}_{\ell=0}^{N-1}$. Then $\{x_\ell\}_{\ell=0}^{N-1}$ can be recovered given $\{X_k\}_{k=0}^{N-1}$ by the inversion formula

$$x_\ell = \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi i k\ell}{N}}. \tag{8.6.1}$$

To see this, we compute $\sum\limits_{k=0}^{N-1}\Big(\sum\limits_{j=0}^{N-1} x_j e^{\frac{-2\pi i kj}{N}}\Big)e^{\frac{2\pi i k\ell}{N}}$ and obtain that

$$\begin{aligned} \sum_{k=0}^{N-1}\Big(\sum_{j=0}^{N-1} x_j e^{\frac{-2\pi i kj}{N}}\Big)e^{\frac{2\pi i k\ell}{N}} &= \sum_{j=0}^{N-1}\Big(x_j \sum_{k=0}^{N-1} e^{\frac{2\pi i k(\ell-j)}{N}}\Big) = Nx_\ell + \sum_{\substack{j=0\\ j\neq\ell}}^{N-1}\Big(x_j \sum_{k=0}^{N-1} e^{\frac{2\pi i k(\ell-j)}{N}}\Big) \\ &= Nx_\ell + \sum_{\substack{j=0\\ j\neq\ell}}^{N-1}\Big(x_j \frac{e^{2\pi i(\ell-j)} - 1}{e^{\frac{2\pi i(\ell-j)}{N}} - 1}\Big) = Nx_\ell. \end{aligned}$$

The map from $\{X_k\}_{k=0}^{N-1}$ to $\{x_\ell\}_{\ell=0}^{N-1}$ is called the ***discrete inverse Fourier transform***.

We note that the inversion formula (8.6.1) is an analogy of

$$f(x) = \sum_{k=-\infty}^{\infty} \widehat{f}_k e^{ikx}$$

for all piecewise constant function f and $x \in \mathbb{R}$ at which f is continuous.

Remark 8.37. Given a sample data $[x_0, x_1, \cdots, x_{N-1}]$ which is the values of a function f on N evenly distributed points on $[0, L)$ (for some unknown $L > 0$), the DFT $[X_0, X_1, \cdots, X_{N-1}]$ also satisfies

$$\begin{aligned} x_\ell &= \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi i k\ell}{N}} = \frac{1}{N}\sum_{k=0}^{[\frac{N-1}{2}]} X_k e^{\frac{2\pi i k\ell}{N}} + \frac{1}{N}\sum_{k=[\frac{N-1}{2}]+1}^{N-1} X_{k-N} e^{\frac{2\pi i(k-N)\ell}{N}} \\ &= \frac{1}{N}\sum_{k=0}^{[\frac{N-1}{2}]} X_k e^{\frac{2\pi i k\ell}{N}} + \frac{1}{N}\sum_{k=-[\frac{N}{2}]}^{-1} X_k e^{\frac{2\pi i k\ell}{N}} = \frac{1}{N}\sum_{k=-[\frac{N}{2}]}^{[\frac{N-1}{2}]} X_k e^{\frac{2\pi i k\ell}{N}}; \end{aligned}$$

thus we may also consider the following approximation:

$$f(x) \approx \frac{1}{N}\sum_{k=-[\frac{N}{2}]}^{[\frac{N-1}{2}]} X_k e^{\frac{2\pi i k x}{L}},$$

where $\approx$ becomes $=$ if $x = \dfrac{L\ell}{N}$, $0 \leqslant \ell \leqslant N-1$. Therefore, comparing with

$$f(x) \approx s_{[\frac{N-1}{2}]}(f, x) = \sum_{k=-[\frac{N-1}{2}]}^{[\frac{N-1}{2}]} \widehat{f}_k e^{\frac{2\pi i k x}{L}},$$

we find that for $0 \leqslant k \leqslant [\dfrac{N-1}{2}]$ each X_k is the coefficient associated with the wave with frequency $\dfrac{k}{L}$. To determine L, we introduce the ***sampling frequency*** F_s which is the number of samples per unit time/length. Then $F_s = \dfrac{N}{L}$ so that X_k is the coefficient associated with the wave with frequency $\dfrac{F_s}{N}k$.

8.6.2 The fast Fourier transform

Let $M = [m_{k\ell}]$ be an $N \times N$ matrix with entry $m_{k\ell}$ defined by

$$m_{k\ell} = e^{\frac{-2\pi i k\ell}{N}} \qquad 0 \leqslant k, \ell \leqslant N-1,$$

and write $\boldsymbol{x} = [x_0, x_1, \cdots, x_{N-1}]^{\mathrm{T}}$ and $\boldsymbol{X} = [X_0, \cdots, X_{N-1}]^{\mathrm{T}}$. Then $\boldsymbol{X} = M\boldsymbol{x}$ and it requires N^2 multiplications to compute $\boldsymbol{X}$. The ***fast Fourier transform***, symbolized by FFT, is a much faster way to compute $\boldsymbol{X}$. In the following, we show that when $N = 2^\gamma$ for some $\gamma \in \mathbb{N}$, then there is a way to compute the DFT with at most $N \log_2 N$ multiplications.

With $N = 2^\gamma$, suppose that $(x_0, \cdots, x_{N-1})$ is a given sequence, and $\{X_k\}_{k=0}^{N-1}$ is the DFT of $\{x_k\}_{k=0}^{N-1}$. Let $\omega = e^{-\frac{2\pi i}{N}}$, and

$$\boldsymbol{x}_{\text{even}} = \begin{bmatrix} x_0 & x_2 & x_4 & \cdots & x_{N-2} \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{x}_{\text{odd}} = \begin{bmatrix} x_1 & x_3 & x_5 & \cdots & x_{N-1} \end{bmatrix}$$

Then

$$\begin{aligned} X_j &= \sum_{\ell=0}^{N-1} x_\ell\, \omega^{j\ell} = \sum_{\substack{0\leqslant \ell\leqslant N-1\\ \ell \text{ is even}}} x_\ell \omega^{j\ell} + \omega^j \sum_{\substack{0\leqslant \ell\leqslant N-1\\ \ell \text{ is odd}}} x_\ell\, \omega^{j(\ell-1)} \\ &= \boldsymbol{x}_{\text{even}} \cdot \begin{bmatrix} \omega^0 & \omega^{2j} & \omega^{4j} & \cdots & \omega^{j(N-2)} \end{bmatrix} + \omega^j \boldsymbol{x}_{\text{odd}} \cdot \begin{bmatrix} \omega^0 & \omega^{2j} & \omega^{4j} & \cdots & \omega^{j(N-2)} \end{bmatrix}. \end{aligned}$$

In particular, for $0 \leqslant j \leqslant \dfrac{N}{2} - 1$,

$$\begin{aligned} X_{\frac{N}{2}+j} &= \boldsymbol{x}_{\text{even}} \cdot \begin{bmatrix} \omega^0 & \omega^{2(\frac{N}{2}+j)} & \omega^{4(\frac{N}{2}+j)} & \cdots & \omega^{(\frac{N}{2}+j)(N-2)} \end{bmatrix} \\ &\quad + \omega^{\frac{N}{2}+j} \boldsymbol{x}_{\text{odd}} \cdot \begin{bmatrix} \omega^0 & \omega^{2(\frac{N}{2}+j)} & \omega^{4(\frac{N}{2}+j)} & \cdots & \omega^{(\frac{N}{2}+j)(N-2)} \end{bmatrix} \\ &= \boldsymbol{x}_{\text{even}} \cdot \begin{bmatrix} \omega^0 & \omega^{2j} & \omega^{4j} & \cdots & \omega^{j(N-2)} \end{bmatrix} - \omega^j \boldsymbol{x}_{\text{odd}} \cdot \begin{bmatrix} \omega^0 & \omega^{2j} & \omega^{4j} & \cdots & \omega^{j(N-2)} \end{bmatrix}, \end{aligned}$$

where we have used the fact that $\omega^{\frac{N}{2}} = -1$ to conclude the equality. We note that

$$\left\{ \boldsymbol{x}_{\text{even}} \cdot \begin{bmatrix} \omega^0 & \omega^{2j} & \omega^{4j} & \cdots & \omega^{j(N-2)} \end{bmatrix} \right\}_{j=0}^{N/2}$$

is exactly the DFT of the sequence $\{x_0, x_2, \cdots, x_{N-2}\}$ and

$$\left\{ \boldsymbol{x}_{\text{odd}} \cdot \begin{bmatrix} \omega^0 & \omega^{2j} & \omega^{4j} & \cdots & \omega^{j(N-1)} \end{bmatrix} \right\}_{j=0}^{N/2}$$

is exactly the DFT of the sequence $\{x_1, x_3, \cdots, x_{N-1}\}$. In other words, to compute the DFT of $\{x_0, \cdots, x_{N-1}\}$, where $N = 2^\gamma$, it suffices to compute the DFTs of the sequence $\{x_0, x_2, \cdots, x_{N-2}\}$ and $\{x_1, x_3, \cdots, x_{N-1}\}$. If the DFTs of the sequences $\{x_0, x_2, \cdots, x_{N-2}\}$ and $\{x_1, x_3, \cdots, x_{N-1}\}$ are known, it requires another $\dfrac{N}{2}$ multiplications to compute the DFT of $\{x_0, x_1, \cdots, x_{N-1}\}$.

Now we compute the total multiplications it requires to compute the DFT of the sequence $\{x_k\}_{k=0}^{2^\gamma-1}$ using the procedure above. Suppose that to compute the DFT of $\{x_k\}_{k=0}^{2^\gamma-1}$ requires $f(\gamma)$ multiplications. Then

$$f(\gamma) = 2f(\gamma - 1) + 2^{\gamma-1}.$$

It is easy to see that it requires no multiplication to compute the DFT of $\{x_0, x_1\}$ since it is simply $\{x_0+x_1, x_0-x_1\}$; thus $f(1)=0$. Solving the iteration relation above, we obtain that $f(\gamma)=2^{\gamma-1}(\gamma-1)$ which implies the total multiplications requires to compute the DFT of $\{x_k\}_{k=0}^{N-1}$, where $N=2^\gamma$, is $\dfrac{N}{2}(\log_2 N-1)$.

Example 8.38. To compute the DFT of $\{x_0, x_1, \cdots, x_7\}$, we first compute the DFT of $\{x_0, x_2, x_4, x_6\}$ and $\{x_1, x_3, x_5, x_7\}$, and it requires another 4 multiplications (to compute the multiplication of ω^j and the j-th term of the DFT of $\{x_1, x_3, x_5, x_7\}$ for $0 \leqslant j \leqslant 3$). Nevertheless, instead of computing the DFT of $\{x_0, x_2, x_4, x_6\}$ and $\{x_1, x_3, x_5, x_7\}$ directly using matrix multiplication $\boldsymbol{X}=M\boldsymbol{x}$, we again divide the sequence of length 4 into further shorter sequence $\{x_0, x_4\}$, $\{x_2, x_6\}$, $\{x_1, x_5\}$ and $\{x_3, x_7\}$. Once the DFT of those sequence of length 2 are computed, it requires another $2 \times 2 = 4$ multiplications to compute the DFT of $\{x_0, x_2, x_4, x_6\}$ and $\{x_1, x_3, x_5, x_7\}$. By Example 8.36, it does not require any multiplications to compute the DFT of sequences of length 2; thus the total multiplications required to compute the DFT of $\{x_0, x_1, \cdots, x_7\}$ is $4+4=8$.

8.7 Fourier Series for Functions of Two Variables

In this section we briefly introduce the Fourier series of complex-valued functions defined on $\Omega \equiv [-L_1, L_1] \times [-L_2, L_2]$. Let

$$L^2(\Omega) = \Big\{ f : \Omega \to \mathbb{C} \,\Big|\, \int_\Omega \big|f(x_1, x_2)\big|^2 \, d(x_2, x_2) < \infty \Big\} \Big/ \sim$$

equipped with the inner product

$$\langle f, g \rangle \equiv \frac{1}{\nu(\Omega)} \int_\Omega f(x_1, x_2) \overline{g(x_1, x_2)} \, d(x_1, x_2),$$

where $\nu(\Omega)$ denotes the area of Ω and $\sim$ again denotes the equivalence relation defined by $f \sim g$ if and only if $f-g=0$ except on a set of measure zero. Let $\mathbf{e}_{k\ell}(\boldsymbol{x}) = e^{i\pi(\frac{k}{L_1}, \frac{\ell}{L_2})\cdot \boldsymbol{x}}$, here $\boldsymbol{x}=(x_1, x_2)$. Then $\{\mathbf{e}_{k\ell}\}_{k,\ell\in\mathbb{Z}}$ is a complete orthonormal set in $L^2(\Omega)$; that is, for each $f \in L^2(\Omega)$, by defining the partial sum

$$s_{n,m}(f, \boldsymbol{x}) = \sum_{k=-n}^{n} \sum_{\ell=-m}^{m} \langle f, \mathbf{e}_{k\ell} \rangle \mathbf{e}_{k\ell}(\boldsymbol{x})$$

we have

$$\lim_{n,m\to\infty} \big\| f - s_{n,m}(f, \cdot) \big\|_{L^2(\Omega)} = 0,$$

where $\|\cdot\|_{L^2(\Omega)}$ is the norm induced by the inner product $\langle\cdot,\cdot\rangle$. The limit of $s_{n,m}(f,\cdot)$, as $n,m\to\infty$, in the inner product space $\big(L^2(\Omega),\langle\cdot,\cdot\rangle\big)$ is denoted by

$$s(f,\cdot)=\sum_{k=-\infty}^{\infty}\sum_{\ell=-\infty}^{\infty}\langle f,\mathbf{e}_{k\ell}\rangle\,\mathbf{e}_{k\ell}$$

and is called the Fourier series of f.

Given a collection of data $\{x_{mn}\}_{0\leqslant n\leqslant M-1,0\leqslant n\leqslant N-1}$, the discrete Fourier transform (or simply DFT) of $\{x_{mn}\}_{0\leqslant n\leqslant M-1,0\leqslant n\leqslant N-1}$ is a double sequence $\{X_{k\ell}\}_{k,\ell\in\mathbb{Z}}$ defined by

$$X_{k\ell}=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}x_{mn}\,\omega_{M}^{mk}\omega_{N}^{n\ell},$$

where $\omega_M=e^{-\frac{2\pi i}{M}}$ and $\omega_N=e^{-\frac{2\pi i}{N}}$. The double sequence $\{X_{k\ell}\}_{k,\ell\in\mathbb{Z}}$ is doubly periodic satisfying $X_{k+M,\ell+N}$ for all $k,\ell\in\mathbb{Z}$; thus we usually only focus on the terms $\{X_{k\ell}\}_{0\leqslant k\leqslant M-1,0\leqslant \ell\leqslant N-1}$. The discrete inverse Fourier transform of a double sequence $\{X_{k\ell}\}_{0\leqslant k\leqslant M-1,0\leqslant \ell\leqslant N-1}$ is a double sequence $\{x_{mn}\}_{m,n\in\mathbb{Z}}$ defined by

$$x_{mn}=\frac{1}{MN}\sum_{k=0}^{M-1}\sum_{\ell=0}^{N-1}X_{k\ell}\,\overline{\omega_{M}}^{mk}\overline{\omega_{N}}^{n\ell},$$

where $\overline{\omega_M}$ and $\overline{\omega_N}$ are complex conjugate of ω_M and ω_N defined above.

8.8 Exercises

§ 8.1 Basic Properties of the Fourier Series

Problem 8.1. 1. Let $f:[-\pi,\pi]$ be a Riemann integrable function. Show that

$$\lim_{k\to\infty}\int_{-\pi}^{\pi}f(x)\cos kx\,dx=\lim_{k\to\infty}\int_{-\pi}^{\pi}f(x)\sin kx\,dx=0\,.$$

2. Show the Riemann-Lebesgue Lemma

> If $f:[-\pi,\pi]\to\mathbb{R}$ is an integrable function, then
> $$\lim_{k\to\infty}\int_{-\pi}^{\pi}f(x)\cos kx\,dx=\lim_{k\to\infty}\int_{-\pi}^{\pi}f(x)\sin kx\,dx=0\,.$$

Hint: First show that for every $\varepsilon>0$ there exists a Riemann integrable function $g:[-\pi,\pi]\to\mathbb{R}$ such that $\displaystyle\int_{-\pi}^{\pi}\big|f(x)-g(x)\big|\,dx<\varepsilon$, then apply the conclusion in 1.

Problem 8.2. Suppose that $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$; that is, f is 2π-periodic Hölder continuous function with exponent α for some $\alpha \in (0,1]$. Show that (without using the Berstein Theorem) the Fourier series of f converges pointwise to f, by completing the following.

1. Explain why it is enough to show that $s_n(f,0) \to f(0)$ as $n \to \infty$. Also explain why we can assume that $f(0) = 0$.

2. Show that
$$\lim_{n\to\infty} \left(s_n(f,0) - \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \frac{\sin nx}{x}\, dx \right) = 0\,.$$
Therefore, it suffices to show that $\displaystyle\lim_{n\to\infty} \int_{-\pi}^{\pi} f(x)\frac{\sin nx}{x}\, dx = 0$ if $f(0) = 0$.

3. Show that if $f \in \mathscr{C}^{0,\alpha}(\mathbb{R})$ and $f(0) = 0$, then the function $y = \dfrac{f(x)}{x}$ is integrable. Apply the Riemann-Lebesgue Lemma to conclude that $s_n(f,0) \to 0$ as $n \to \infty$.

Problem 8.3. Assuming that the improper integral $\displaystyle\int_0^\infty \frac{\sin x}{x}\, dx = I$ exists. Establish its value by first using the Riemann-Lebesgue lemma to show that
$$I = \lim_{n\to\infty} \int_0^{n\pi} \frac{\sin x}{x}\, dx = \frac{\pi}{2} \lim_{n\to\infty} \int_{-\pi}^{\pi} D_n(x)\, dx\,,$$
where D_n is the Dirchlet kernel.

§ 8.2 Uniform Convergence of the Fourier Series

Problem 8.4. Let f be the 2π-periodic function defined by $f(x) = \cosh(x) = \dfrac{e^x + e^{-x}}{2}$ for $|x| \leqslant \pi$. Express it as a Fourier series and compute
$$\sum_{k=0}^{\infty} \frac{1}{k^2+1}\,.$$

Problem 8.5. Let f be the 2π-periodic function defined by $f(x) = \cos(ax)$ for $|x| \leqslant \pi$, where a is not an integer. Express it as a Fourier series and deduce the identity
$$\pi\cot(\pi a) = \lim_{N\to\infty} \sum_{n=-N}^{N} \frac{1}{a+n} \qquad \forall\, a \neq \mathbb{Z}\,.$$

Problem 8.6. Let $\alpha > \dfrac{1}{\pi}$ and $f : \mathbb{R} \to \mathbb{R}$ be an even 2π-periodic function whose value on $[0,\pi]$ is given by
$$f(x) = \begin{cases} \alpha - \alpha^2 x & \text{if } x \in [0, 1/\alpha], \\ 0 & \text{if } x \in (1/\alpha, \pi]. \end{cases}$$

1. Find the Fourier coefficients $\{c_k\}_{k=0}^{\infty}$, $\{s_k\}_{k=1}^{\infty}$ and the Fourier series of f.

2. Explain why the Fourier series of f converges uniformly to f on $\mathbb{R}$.

3. Use the Fourier series of f to show that $\displaystyle\sum_{k=1}^{\infty} \frac{\cos^2 k}{k^2} = \frac{1}{4} + \sum_{k=1}^{\infty} \frac{\cos k}{k^2}$.

4. For each α, denote the Fourier coefficients of f by $\{c_k(\alpha)\}_{k=0}^{\infty}$, $\{s_k(\alpha)\}_{k=1}^{\infty}$ (and note that $s_k(\alpha) = 0$ for all $k \in \mathbb{N}$). Prove that $\lim\limits_{\alpha\to\infty} c_k(\alpha)$ exists and the limit is independent of k.

Problem 8.7. For each $r \in [0, 1)$, define $f_r : \mathbb{R} \to \mathbb{R}$ by

$$f_r(t) = \ln(1 + r^2 - 2r \cos t)\,.$$

Clearly f_r is a 2π-periodic even function. Show that the Fourier series of f_r is given by

$$-2 \sum_{k=1}^{\infty} \frac{r^k}{k} \cos kt$$

by the following steps.

1. For $z \in \mathbb{C}$ and $|z| < 1$, we have $\dfrac{1}{1-z} = \sum\limits_{k=0}^{\infty} z^k$. Let $z = re^{it}$ in this sum and conclude that
$$\frac{1}{1 - re^{it}} = \sum_{k=0}^{\infty} r^k(\cos kt + i \sin kt)\,.$$
Take the imaginary part of the formula above to conclude that
$$\sum_{k=1}^{\infty} r^k \sin kt = \frac{r \sin t}{1 + r^2 - 2r \cos t}\,.$$

2. Since the sum on the left-hand side converges uniformly if $r \in [0, 1)$, for a fixed $r \in [0, 1)$ we can integrate in t term-by-term to conclude that
$$\sum_{k=1}^{\infty} \frac{r^k}{k} \cos kt = -\frac{1}{2} \ln(1 + r^2 - 2r \cos t) + C\,.$$
Determine this constant C.

3. Find $\sum\limits_{k=1}^{\infty} \dfrac{1}{k2^k}$ and $\sum\limits_{k=1}^{\infty} \dfrac{(-1)^k}{k3^k}$.

4. Show that $\sum\limits_{k=1}^{\infty} \dfrac{(-1)^{k+1} r^k}{k} = \ln(1+r)$ for $r \in (-1,1)$.

Problem 8.8. Prove Lemma 8.15.

Problem 8.9. Let f be a 2π-periodic Lipchitz function; that is, $f \in \mathscr{C}^{0,1}(\mathbb{T})$. Show that for $n \geqslant 2$,

$$\|f - F_{n-1} \star f\|_\infty \leqslant \frac{1 + 2\log n}{2n} \pi \|f\|_{\mathscr{C}^{0,1}(\mathbb{T})} \tag{8.8.1}$$

and

$$\big\|f - s_n(f,\cdot)\big\|_\infty \leqslant \frac{2\pi(1+\log n)^2}{n} \|f\|_{\mathscr{C}^{0,1}(\mathbb{T})}\,. \tag{8.8.2}$$

Inequality (8.8.2) provides the rate of convergence of the Fourier series to Lipschitz functions. What is the rate of convergence if $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$?

Hint: For (8.8.1), apply the estimate

$$F_n(x) \leqslant \min\left\{\frac{n+1}{2\pi}, \frac{\pi}{2(n+1)x^2}\right\}$$

in the following inequality:

$$\big|f(x) - F_{n-1} \star f(x)\big| \leqslant \Big[\int_{-\delta}^{\delta} + \int_{-\pi}^{-\delta} + \int_{\delta}^{\pi}\Big] \big|f(x) - f(x-y)\big| F_{n-1}(y)\,dy$$

with $\delta = \dfrac{\pi}{n}$. For (8.8.2), use (8.2.7) and note that

$$\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty \leqslant \|f - F_n \star f\|_\infty\,.$$

Problem 8.10. In this problem, we are concerned with the following

Theorem 8.39 (Bernstein)**.** *Suppose that f is a 2π-periodic function such that for some constant C and $\alpha \in (0,1)$,*

$$\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_\infty \leqslant C n^{-\alpha}$$

for all $n \in \mathbb{N}$. Then $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$.

Complete the following to prove the theorem.

1. Show that

$$\|p'\|_\infty \leqslant n\|p\|_\infty \qquad \forall\, p \in \mathscr{P}_n(\mathbb{T}). \tag{8.8.3}$$

2. Choose $p_n \in \mathscr{P}_n(\mathbb{T})$ such that $\|f - p_n\|_\infty \leqslant 2Cn^{-\alpha}$ for $n \in \mathbb{N}$. Define $q_0 = p_1$, and $q_n = p_{2^n} - p_{2^{n-1}}$ for $n \in \mathbb{N}$.

   (a) Show that $\sum\limits_{n=0}^{\infty} q_n = f$ and the convergence is uniform.

   (b) Show that
   $$\big|q_n(x) - q_n(y)\big| \leqslant 6Cn2^{n(1-\alpha)}|x-y| \quad \text{and} \quad \big|q_n(x) - q_n(y)\big| \leqslant 12C2^{-n\alpha}.$$

   (c) For any $x, y \in \mathbb{T}$ with $|x-y| \leqslant 1$, choose $m \in \mathbb{N}$ such that $2^{-m} \leqslant |x-y| \leqslant 2^{1-m}$. Then use the inequality
   $$\big|f(x) - f(y)\big| \leqslant \sum_{n=0}^{m-1} \big|q_n(x) - q_n(y)\big| + \sum_{n=m}^{\infty} \big|q_n(x) - q_n(y)\big|$$
   to show that $\big|f(x) - f(y)\big| \leqslant B|x-y|^\alpha$ for some constant $B > 0$.

Hint of 1: Suppose the contrary that there exists $p \in \mathscr{P}_n(\mathbb{T})$ such that $\|p'\|_\infty > n\|p\|_\infty$. By rescaling p and relabeling points in $\mathbb{T}$ if necessary, without loss of generality we can assume that
$$\|p'\|_\infty > n\,, \quad \|p\|_\infty < 1\,, \quad \text{and} \quad p'(0) = \|p'\|_\infty\,.$$
Choose $\gamma \in \big[-\frac{\pi}{2n}, \frac{\pi}{2n}\big]$ such that $\sin(n\gamma) = -p(0)$, and define $r(x) = \sin n(x-\gamma) - p(x)$. Show that r has at least $2n+2$ distinct zeros in $[\alpha_{-n}, \alpha_n] \equiv \big[-\pi + \gamma + \frac{\pi}{2n}, \pi + \gamma + \frac{\pi}{2n}\big]$ by showing that r has at least one zero in (α_k, α_{k+1}), where $\alpha_k = \gamma + \frac{\pi}{n}\big(k + \frac{1}{2}\big)$ for each $|k| \leqslant n$, while r has at least 3 distinct zeros in (α_s, α_{s+1}) if $0 \in (\alpha_s, \alpha_{s+1})$ (in fact, $s = -1$). On the other hand, the fact that $r \in \mathscr{P}_n(\mathbb{T})$ implies that r has at most $2n$ distinct zeros in $\mathbb{T}$ unless r is the zero function which leads to a contradiction.

§ 8.3 Cesàro Mean of Fourier Series

Problem 8.11. Let $f \in \mathscr{C}(\mathbb{T})$ and $\{\widehat{f}_k\}_{k=-\infty}^{\infty}$ be the Fourier coefficients defined in Remark 8.2. Show that if $\sum\limits_{k=-\infty}^{\infty} \big|\widehat{f}_k\big| < \infty$, then the Fourier series of f converges uniformly to f on $\mathbb{R}$.

Problem 8.12. A family of functions $\big\{\varphi_n \in \mathscr{C}(\mathbb{T}) \,\big|\, n \in \mathbb{N}\big\}$ is called an ***approximation of the identity*** if

(1) $\varphi_n(x) \geqslant 0$; (2) $\displaystyle\int_{\mathbb{T}} \varphi_n(x)\,dx = 1$ for every $n \in \mathbb{N}$;

(3) $\displaystyle\lim_{n\to\infty} \int_{\delta \leqslant |x| \leqslant \pi} \varphi_n(x)\,dx = 0$ for every $\delta > 0$, here we identify $\mathbb{T}$ with the interval $[-\pi, \pi]$.

Show that if $\{\varphi_n\}_{n=1}^{\infty}$ is an approximation of the identity and $f \in \mathscr{C}(\mathbb{T})$, then $\{\varphi_n \star f\}_{n=1}^{\infty}$ converges uniformly to f as $n \to \infty$.

Problem 8.13. In this problem we show that the collection of trigonometric polynomials $\mathscr{P}(\mathbb{T})$ (defined in Corollary 7.65) is dense in $\mathscr{C}(\mathbb{T})$ in another way. Complete the following.

1. Let $\varphi_n(x) = c_n(1 + \cos x)^n$, where c_n is chosen so that $\displaystyle\int_{\mathbb{T}} \varphi_n(x)\,dx = 1$. Show that
$$c_n = \frac{2^{n-1}}{\pi}\frac{(n!)^2}{(2n)!}\,.$$

2. Show that for each $0 < \delta < \pi$,
$$\lim_{n\to\infty} \int_{\delta \leqslant |x| \leqslant \pi} \varphi_n(x)\,dx = 0\,.$$
In other words, $\{\varphi_n\}_{n=1}^{\infty}$ is an approximation of the identity. Therefore, Problem 8.12 shows that $\{\varphi_n \star f\}_{n=1}^{\infty}$ converges uniformly to f as $n \to \infty$ if $f \in \mathscr{C}(\mathbb{T})$.

3. Show that $\mathscr{P}(\mathbb{T})$ is dense in $\mathscr{C}(\mathbb{T})$.

§ 8.4 Convergence of Fourier Series for Functions with Jump Discontinuity

Problem 8.14. Compute the Fourier series of the function $f : (-\pi, \pi) \to \mathbb{R}$ defined by
$$f(x) = \left\{\begin{array}{cl} 0 & -\pi < x < 0\,, \\ \pi - x & 0 \leqslant x < \pi\,, \end{array}\right.$$
and show that
$$1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \sum_{k=1}^{\infty} \frac{1}{(2k-1)^2} = \frac{\pi^2}{8}\,. \tag{8.8.4}$$
Also use the Fourier series of the function $y = x^2$
$$s(x^2, x) = \frac{\pi^2}{3} + 4\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2}\cos kx$$
to conclude (8.8.4).

Problem 8.15. The proof of Theorem 8.25 only establishes the validity of the theorem for the case $L = \pi$. Use this fact to show that Theorem 8.25 also holds for general $L > 0$.

Problem 8.16. For a given function $f : [0, L] \to \mathbb{R}$, the even extension of f is a function $\bar{f} : [-L, L] \to \mathbb{R}$ such that

$$\bar{f}(x) = f(x) \text{ if } x \in [0, L) \qquad \text{and} \qquad \bar{f}(x) = f(-x) \text{if } x \in [-L, 0)\,.$$

1. Let $f : [0, L] \to \mathbb{R}$ be an integrable function. The cosine series of f is the Fourier series of the even extension of f. Find the cosine series of f.

2. Suppose in addition $f : [0, L] \to \mathbb{R}$ is piecewise Hölder continuous with exponent $\alpha \in (0, 1]$. Show that the cosine series of f at $x_0 \in (0, L)$ converges to $\dfrac{f(x_0^+) + f(x_0^-)}{2}$.

Problem 8.17. For a given function $f : [0, L] \to \mathbb{R}$, the odd extension of f is a function $\bar{f} : [-L, L] \to \mathbb{R}$ such that

$$\bar{f}(x) = f(x) \text{ if } x \in [0, L) \qquad \text{and} \qquad \bar{f}(x) = -f(-x) \text{if } x \in [-L, 0)\,.$$

1. Let $f : [0, L] \to \mathbb{R}$ be an integrable function. The sine series of f is the Fourier series of the odd extension of f. Find the cosine series of f.

2. Suppose in addition $f : [0, L] \to \mathbb{R}$ is piecewise Hölder continuous with exponent $\alpha \in (0, 1]$. Show that the sine series of f at $x_0 \in (0, L)$ converges to $\dfrac{f(x_0^+) + f(x_0^-)}{2}$.

Problem 8.18. Let f be the sinc function defined by

$$f(x) = \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0\,, \\ 1 & \text{if } x = 0\,. \end{cases}$$

Show that

$$f(x) = \frac{b_0}{2} + \sum_{k=1}^{\infty} b_n \cos(nx)\,, \qquad \forall\, x \in \mathbb{R}\,,$$

where $b_n = \dfrac{1}{\pi} \displaystyle\int_{(n-1)\pi}^{(n+1)\pi} \frac{\sin x}{x}\, dx$. Use this result to compute $\displaystyle\int_0^{\infty} \frac{\sin x}{x}\, dx$.

§ 8.5 The Inner-Product Point of View

Problem 8.19. Let $f : \mathbb{R} \to \mathbb{R}$ be the function given in Problem 8.6. Find $\sum\limits_{k=1}^{\infty} \dfrac{\sin^4 k}{k^4}$ using the Parseval identity corresponding to f.

Problem 8.20. Use the Fourier series of the function $f : (-\pi, \pi) \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & -\pi < x < 0\,, \\ \pi - x & 0 \leqslant x < \pi\,, \end{cases}$$

and compute

$$1 + \frac{1}{2^4} + \frac{1}{3^4} + \frac{1}{4^4} + \cdots = \sum_{k=1}^{\infty} \frac{1}{k^4}\,.$$

Problem 8.21. Use the Fourier series of the function $f : [-\pi, \pi] \to \mathbb{R}$ given by $f(x) = x^3 - \pi^2 x$ to find the values of $\sum\limits_{k=1}^{\infty} \dfrac{(-1)^k}{(2k-1)^3}$ and $\sum\limits_{k=1}^{\infty} \dfrac{1}{k^6}$.

Problem 8.22. For each $n \in \mathbb{Z}$, define the Bessel functions $J_n(x)$ through the Fourier series by

$$e^{ix \sin t} = \sum_{n=-\infty}^{\infty} J_n(x) e^{int}\,.$$

Compute $\sum\limits_{n=-\infty}^{\infty} |J_n(x)|^2$ for $x \in \mathbb{R}$.

Problem 8.23. Let $f : [-L, L] \to \mathbb{R}$ be square integrable; that is, f is Riemann measurable and $\displaystyle\int_{-L}^{L} f(x)^2\, dx < \infty$. Show that if $\{c_k\}_{k=0}^{\infty}$ and $\{s_k\}_{k=1}^{\infty}$ be the Fourier coefficient of f, then

$$\frac{1}{2L} \int_{-L}^{L} f(x)^2\, dx = \frac{c_0^4}{4} + \frac{1}{2} \sum_{k=1}^{\infty} (c_k^2 + s_k^2)\,.$$

Problem 8.24. Let $f : [0, L] \to \mathbb{R}$ be a square integrable function.

1. Suppose that $\dfrac{c_0}{2} + \sum\limits_{k=1}^{\infty} c_k \cos \dfrac{k\pi x}{L}$ is the cosine series of f. Find $\sum\limits_{k=1}^{\infty} c_k^2$ in terms of integrals of f and f^2.
2. Suppose that $\sum\limits_{k=1}^{\infty} s_k \sin \dfrac{k\pi x}{L}$ is the sine series of f. Find $\sum\limits_{k=1}^{\infty} s_k^2$ in terms of integral of f^2.

Problem 8.25. Expand the function $\cos x$ as a sine series on the interval $(0,\pi)$. Use the result to compute

$$\sum_{n=1}^{\infty} \frac{n^2}{(4n^2-1)^2}\,.$$

How about expanding $\cos x$ as a sine series on the interval $(0,\pi/2)$?

Problem 8.26. This problem contributes to another proof of showing that the Fourier series of f converges uniformly to f on $\mathbb{R}$ if $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ for $\frac{1}{2} < \alpha \leqslant 1$. Complete the following.

1. Let $f : \mathbb{R} \to \mathbb{R}$ be 2π-periodic such that f is Riemann integrable on $[-\pi,\pi]$. Show that

$$\widehat{f}_k = -\frac{1}{2\pi}\int_{-\pi}^{\pi} f\big(x+\frac{\pi}{k}\big)e^{-ikx}\,dx$$

and hence

$$\widehat{f}_k = \frac{1}{4\pi}\int_{-\pi}^{\pi} \big[f(x)-f\big(x+\frac{\pi}{k}\big)\big]e^{-ikx}\,dx\,.$$

Therefore, if $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, the Fourier coefficients $\widehat{f}_k$ satisfies $|\widehat{f}_k| \leqslant \dfrac{\pi^\alpha \|f\|_{\mathscr{C}^{0,\alpha}(\mathbb{T})}}{2k^\alpha}$.

2. Let $f : \mathbb{R} \to \mathbb{R}$ be 2π-periodic such that f is Riemann integrable on $[-\pi,\pi]$. Show that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \big|f(x+h)-f(x-h)\big|^2\,dx = \sum_{k=-\infty}^{\infty} 4\sin^2(kh)|\widehat{f}_k|^2\,.$$

Therefore, if $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, the Fourier coefficients $\widehat{f}_k$ satisfies

$$\sum_{k=-\infty}^{\infty} \sin^2(kh)|\widehat{f}_k|^2 \leqslant \|f\|^2_{\mathscr{C}^{0,\alpha}(\mathbb{T})} 2^{2(\alpha-1)}|h|^{2\alpha} \tag{8.8.5}$$

3. Let $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, and $p \in \mathbb{N}$. Show that

$$\sum_{2^{p-1}\leqslant|k|<2^p} |\widehat{f}_k|^2 \leqslant \frac{\|f\|^2_{\mathscr{C}^{0,\alpha}(\mathbb{T})}\pi^{2\alpha}}{2^{2\alpha p+1}}\,.$$

Hint: Let $h = \dfrac{\pi}{2^{p+1}}$ in (8.8.5).

4. Show that if $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ for some $\frac{1}{2} < \alpha \leqslant 1$, then $\sum\limits_{k=-\infty}^{\infty} |\widehat{f}_k| < \infty$; thus Problem 8.11 implies that the Fourier series of f converges uniformly to f on $\mathbb{R}$.

§ 8.6 The Discrete Fourier "Transform" and the Fast Fourier "Transform"

Problem 8.27. Determine the discrete Fourier transform of the sequence $(1, 0, 0, -1)$. Check the inversion formula.

Problem 8.28. Determine the discrete Fourier transform of the N-periodic sequence $x_n = \sin\dfrac{n\pi}{N}$, $n = 0, \cdots, N-1$.

Problem 8.29. Give a version of the Parseval identity for discrete Fourier transform.

Problem 8.30. Let $x(n)$ be N-periodic, and

$$x(n) = \begin{cases} 1 & \text{if } 0 \leqslant n \leqslant k-1\,, \\ 0 & \text{if } k \leqslant n \leqslant N-1\,. \end{cases}$$

Compute the discrete Fourier transform. Using the Parseval identity (for DFT), compute the sum

$$\sum_{n=1}^{N-1} \frac{1-\cos\frac{2\pi nk}{N}}{1-\cos\frac{2\pi n}{N}}\,.$$

Chapter 9

Fourier Transforms

Before introducing the Fourier transform, let us "motivate" the idea a little bit. In Section 8.5 we show that $\{\mathbf{e}_n\}_{n=-\infty}^{\infty}$, where $\mathbf{e}_n(x) = e^{inx}$, is a complete orthonormal set in $L^2(\mathbb{T})$. Similarly, let $L^2([-K,K])$ denote the inner-product space

$$L^2([-K,K]) = \big\{f : [-K,K] \to \mathbb{C} \,\big|\, f \text{ is square integrable}\big\}\big/ \sim$$

equipped with the inner product

$$\langle f, g\rangle = \frac{1}{2K}\int_{-K}^{K} f(x)\overline{g(x)}\,dx\,,$$

where $\sim$ denotes the equivalence relation $f \sim g$ if and only if $f - g = 0$ except on a set of measure zero. Then the set $\Big\{\exp\big(\frac{in\pi x}{K}\big)\Big\}_{n=-\infty}^{\infty}$ is a complete orthonormal set in $L^2([-K,K])$; that is, any functions $f \in L^2([-K,K])$ can be expressed as

$$f(x) = \sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{\frac{in\pi x}{K}}\,, \quad \text{where} \quad \widehat{f}(n) = \frac{1}{2K}\int_{-K}^{K} f(y)e^{-\frac{in\pi y}{K}}\,dy\,. \tag{9.0.1}$$

Moreover, $\sum\limits_{n=-\infty}^{\infty} |\widehat{f}(n)|^2 = \dfrac{1}{2K}\displaystyle\int_{-K}^{K} |f(x)|^2\,dx$. In other words, there is a one-to-one correspondence between $f \in L^2([-K,K])$ and $\widehat{f} \in \ell_2$, where ℓ^2 is the collection of square summable sequences; that is,

$$\ell^2 = \Big\{\{a_n\}_{n=-\infty}^{\infty}\,\Big|\, \sum_{n=-\infty}^{\infty} |a_n|^2 < \infty\Big\}.$$

We look for a space X so that there is also a one-to-one correspondence between the square integrable functions on $\mathbb{R}$ and X. Intuitively, we can check what "might" happen by letting $K \to \infty$ in (9.0.1).

Suppose that $K \gg 1$ and $K \in \mathbb{N}$. Making use of the Riemann sum to approximate the integral (by partition $[-K\pi, K\pi]$ into $2K^2$ intervals), we find that

$$\begin{aligned}
f(x) &= \frac{1}{2K}\sum_{n=-\infty}^{\infty}\int_{-K}^{K} f(y)e^{\frac{in\pi(x-y)}{K}}\,dy \approx \frac{1}{2K}\sum_{n=-K^2}^{K^2-1}\int_{-K}^{K} f(y)\exp\Big[i\frac{n\pi}{K}(x-y)\Big]dy \\
&= \frac{1}{2\pi}\int_{-K}^{K}\Big(\sum_{n=-K^2}^{K^2-1} f(y)\exp\Big[i\frac{n\pi}{K}(x-y)\Big]\frac{\pi}{K}\Big)\,dy \\
&\approx \frac{1}{2\pi}\int_{-K}^{K}\Big(\int_{-K\pi}^{K\pi} f(y)\exp\big(i\xi(x-y)\big)\,d\xi\Big)\,dy = \frac{1}{2\pi}\int_{-K\pi}^{K\pi}\Big(\int_{-K}^{K} f(y)e^{i\xi(x-y)}\,dy\Big)\,d\xi \\
&\approx \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\Big[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(y)e^{-i\xi y}dy\Big]e^{i\xi x}d\xi\,.
\end{aligned}$$

Therefore, if we define $\widehat{f}(\xi) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{\mathbb{R}} f(y)e^{-iy\xi}\,dy$, then the formal computation above suggests that

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \widehat{f}(\xi)e^{i\xi x}d\xi\,. \tag{9.0.2}$$

In the rest of this chapter, we are going to verify the identity above rigorously (for functions f with certain properties).

9.1 The Definition of the Fourier Transform

For notational convenience, we **abuse** the following notion from real analysis.

Definition 9.1. The space $L^1(\mathbb{R}^n)$ consists of all complex-valued functions that are integrable on $\mathbb{R}^n$ and whose integrals are absolutely convergent. In other words,

$$L^1(\mathbb{R}^n) = \Big\{f : \mathbb{R}^n \to \mathbb{C}\,\Big|\, \int_{\mathbb{R}^n} |f(x)|\,dx < \infty\Big\};$$

that is, $f \in L^1(\mathbb{R}^n)$ if the limit $\displaystyle\lim_{R\to\infty}\int_{B(0,R)} \big|f(x)\big|\,dx = \|f\|_{L^1(\mathbb{R}^n)}$ exists.

Remark 9.2. Even though we have not defined the integral for complex-valued function, the definition of $L^1(\mathbb{R}^n)$ should be clear: when f is complex-valued function, the absolute integrability of f is equivalent to that the real part and the imaginary part of f are both absolutely integrable, and

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\,dx &= \int_{\mathbb{R}^n} \mathrm{Re}(f)(x)\,dx + i\int_{\mathbb{R}^n} \mathrm{Im}(f)(x)\,dx \\ &= \int_{\mathbb{R}^n} \frac{f(x)+\overline{f(x)}}{2}\,dx + \int_{\mathbb{R}^n} \frac{f(x)-\overline{f(x)}}{2}\,dx\,,\end{aligned}$$

where $\overline{f(x)}$ is the complex conjugate of $f(x)$.

Definition 9.3. For all $f \in L^1(\mathbb{R}^n)$, the Fourier transform of f, denoted by $\mathscr{F}f$ or $\widehat{f}$, is a function defined by

$$(\mathscr{F}f)(\xi) = \widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(x)e^{-ix\cdot\xi}dx \qquad \forall\,\xi \in \mathbb{R}^n,$$

where $x \cdot \xi = x_1\xi_1 + x_2\xi_2 + \cdots + x_n\xi_n$.

9.2 Some Properties of the Fourier Transform

Proposition 9.4. *$\mathscr{F} : L^1(\mathbb{R}^n) \to \mathscr{C}_b(\mathbb{R}^n;\mathbb{C})$, and*

$$\|\mathscr{F}f\|_\infty \equiv \sup_{\xi\in\mathbb{R}^n}\big|(\mathscr{F}f)(\xi)\big| \leqslant \|f\|_{L^1(\mathbb{R}^n)}\,. \tag{9.2.1}$$

Proof. Let $\xi \in \mathbb{R}^n$, and $\{\xi_k\}_{k=1}^\infty$ be a sequence converging to ξ. Define

$$g_k(x) = \frac{1}{\sqrt{2\pi}^n} f(x)e^{ix\cdot\xi_k} \quad \text{and} \quad g(x) = \frac{1}{\sqrt{2\pi}^n} f(x)e^{ix\cdot\xi}.$$

Then $\{g_k(x)\}_{k=1}^\infty$ converges to $g(x)$ for all $\big\{x \in \mathbb{R}^n \,\big|\, \big|g(x)\big| < \infty\big\}$, and for each $k \in \mathbb{N}$, g_k is integrable and $|g_k| \leqslant |f|$; thus the Dominated Convergence Theorem (Theorem 6.102) implies that

$$(\mathscr{F}f)(\xi) = \int_{\mathbb{R}^n} g(x)\,dx = \lim_{k\to\infty}\int_{\mathbb{R}^n} g_k(x)\,dx = \lim_{k\to\infty}(\mathscr{F}f)(\xi_k)\,.$$

Therefore, $(\mathscr{F}f)$ is continuous on $\mathbb{R}^n$. The validity of (9.2.1) should be clear, and is left as an exercise. ◻

Definition 9.5. A function f on $\mathbb{R}^n$ is said to have rapid decrease/decay if for all integers $N \geqslant 0$, there exists a_N such that

$$|x|^N|f(x)| \leqslant a_N \quad \text{as} \quad x \to \infty\,.$$

Definition 9.6. The *Schwartz space* $\mathscr{S}(\mathbb{R}^n)$ is the collection of all (complex-valued) smooth functions f on $\mathbb{R}^n$ such that f and all of its derivatives have rapid decrease. In other words,

$$\mathscr{S}(\mathbb{R}^n) = \big\{u \in \mathscr{C}^\infty(\mathbb{R}^n) \,\big|\, |\cdot|^N D^k u \text{ is bounded for all } k, N \in \mathbb{N} \cup \{0\}\big\}.$$

Elements in $\mathscr{S}(\mathbb{R}^n)$ are called Schwartz functions.

The reader is encouraged to verify the following basic properties of $\mathscr{S}(\mathbb{R}^n)$:

1. $\mathscr{S}(\mathbb{R}^n)$ is a vector space.
2. $\mathscr{S}(\mathbb{R}^n)$ is an algebra under the pointwise product of functions.
3. $pu \in \mathscr{S}(\mathbb{R}^n)$ for all $u \in \mathscr{S}(\mathbb{R}^n)$ and all polynomial functions p.
4. $\mathscr{S}(\mathbb{R}^n)$ is closed under differentiation.
5. $\mathscr{S}(\mathbb{R}^n)$ is closed under translations and multiplication by complex exponentials $e^{ix\cdot\xi}$.

Remark 9.7. Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and $\mathscr{C}_c^\infty(\Omega)$ denote the collection of all smooth functions with compact support in Ω (or equivalently, compactly supported in Ω); that is,

$$\mathscr{C}_c^\infty(\Omega) \equiv \big\{u \in \mathscr{C}^\infty(\Omega) \,\big|\, \{x \in \Omega \,|\, f(x) \neq 0\} \subset\subset \Omega\big\},$$

then $\mathscr{C}_c^\infty(\mathbb{R}^n) \subseteq \mathscr{S}(\mathbb{R}^n)$. The set $\mathrm{cl}\big(\big\{x \in \Omega \,\big|\, f(x) \neq 0\big\}\big)$, where the closure is taken in the metric space $(\Omega, |\cdot|)$, is called the ***support*** of f and is denoted by $\mathrm{supp}(f)$.

The prototype element of $\mathscr{S}(\mathbb{R}^n)$ is $e^{-|x|^2}$ which is not compactly supported, but has rapidly decreasing derivatives.

The following lemma allows us to take the Fourier transform of Schwartz functions.

Lemma 9.8. *If $f \in \mathscr{S}(\mathbb{R}^n)$, then $f \in L^1(\mathbb{R}^n)$.*

Proof. If $f \in \mathscr{S}(\mathbb{R}^n)$, then $(1+|x|)^{n+1}|f(x)| \leqslant C$ for some $C > 0$. Therefore, with ω_{n-1} denoting the the surface area of the $(n-1)$-dimensional unit sphere,

$$\begin{aligned}\int_{\mathbb{R}^n} \big|f(x)\big|\,dx &\leqslant \int_{\mathbb{R}^n} \frac{C}{(1+|x|)^{n+1}}dx = \int_{\mathbb{S}^{n-1}}\int_0^\infty \frac{C}{(1+r)^{n+1}} r^{n-1} dr dS \\ &\leqslant C\omega_{n-1}\int_0^\infty (1+r)^{-2}dr = C\omega_{n-1}\end{aligned}$$

which is a finite number. □

Next we show that $\widehat{f}$ is differentiable if $f \in \mathscr{S}(\mathbb{R}^n)$. Note that if $f \in \mathscr{S}(\mathbb{R}^n)$, then the function $y_j = x_j f(x)$ belongs to $\mathscr{S}(\mathbb{R}^n)$ for all $1 \leqslant j \leqslant n$.

Lemma 9.9. *If $f \in \mathscr{S}(\mathbb{R}^n)$, then $\widehat{f}$ is differentiable, and for each $j \in \{1, \cdots, n\}$, $\dfrac{\partial \widehat{f}}{\partial \xi_j}$ is given by*

$$\frac{\partial \widehat{f}}{\partial \xi_j}(\xi) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} (-ix_j) f(x) e^{-ix\cdot\xi} dx = \mathscr{F}_x\Big[\frac{1}{i} x_j f(x)\Big](\xi). \tag{9.2.2}$$

Proof. Let $\xi \in \mathbb{R}^n$ and $1 \leqslant j \leqslant n$ be given, and $\{h_k\}_{k=1}^\infty$ be a non-zero sequence converging to 0. Define

$$g_k(x) = \frac{1}{\sqrt{2\pi}^n} f(x) \frac{e^{ix\cdot(\xi + h_k \mathbf{e}_j)} - e^{ix\cdot\xi}}{h_k} = \frac{1}{\sqrt{2\pi}^n} f(x) e^{-ix\cdot\xi} \frac{e^{-ix_j h_k} - 1}{h_k}.$$

Note that the mean value theorem implies that

$$\Big|\frac{e^{-ix_j h_k} - 1}{h_k}\Big| = \Big|\frac{\cos(x_j h_k) - \cos 0}{h_k} - i\frac{\sin(x_j h_k)}{h_k}\Big| \leqslant 2|x_j|;$$

thus

$$\big|g_k(x)\big| \leqslant \frac{2}{\sqrt{2\pi}^n}\big|x_j f(x)\big| \qquad \forall\, x \in \mathbb{R}^n \text{ and } k \in \mathbb{N}.$$

Moreover, the fact that $f \in \mathscr{S}(\mathbb{R}^n)$ implies that the function $y = \dfrac{2}{\sqrt{2\pi}^n}\big|x_j f(x)\big|$ is integrable on $\mathbb{R}^n$. Therefore, by the fact that

$$\int_{\mathbb{R}^n} g_k(x)\, dx = \frac{\widehat{f}(\xi + h_j \mathbf{e}_k) - \widehat{f}(\xi)}{h_j} \qquad \text{and} \qquad \lim_{k\to\infty} g_k(x) = \frac{1}{\sqrt{2\pi}^n}(-ix_j) f(x) e^{-ix\cdot\xi},$$

we conclude from the Dominated Convergence Theorem (Theorem 6.102) that

$$\lim_{k\to\infty} \frac{\widehat{f}(\xi + h_j \mathbf{e}_k) - \widehat{f}(\xi)}{h_j} = \lim_{k\to\infty} \int_{\mathbb{R}^n} g_k(x)\, dx = \int_{\mathbb{R}^n} \frac{1}{\sqrt{2\pi}^n}(-ix_j) f(x) e^{-ix\cdot\xi}\, dx. \qquad \square$$

Corollary 9.10. *If $f \in \mathscr{S}(\mathbb{R}^n)$, then $\widehat{f} \in \mathscr{C}^\infty(\mathbb{R}^n)$. Moreover, if $\alpha = (\alpha_1, \cdots, \alpha_n)$ is a multi-index,*

$$D_\xi^\alpha \widehat{f}(\xi) = \frac{1}{i^{|\alpha|}} \mathscr{F}_x\Big[x_1^{\alpha_1} \cdots x_n^{\alpha_n} f(x)\Big](\xi).$$

Lemma 9.11. *If $f \in \mathscr{S}(\mathbb{R}^n)$, then for $j \in \{1, 2, \cdots, n\}$, $\mathscr{F}_x\Big[\dfrac{1}{i}\dfrac{\partial f}{\partial x_j}(x)\Big](\xi) = \xi_j \widehat{f}(\xi)$.*

Proof. W.L.O.G., we assume that $j = n$. Write $x = (x', x_n)$. Since $f \in \mathscr{S}(\mathbb{R}^n)$, there exists $C > 0$ such that

$$(1+|x'|)^n |x_n| \big|f(x', x_n)\big| \leqslant C \qquad \forall\, x = (x', x_n) \in \mathbb{R}^n .$$

Then

1. For each $x' \in \mathbb{R}^{n-1}$, $f(x', \pm R) \to 0$ as $R \to \infty$.
2. The function $g : \mathbb{R}^{n-1} \to \mathbb{R}$ defined by $g(x') = \dfrac{1}{(1+|x'|)^n}$ is integrable on $\mathbb{R}^{n-1}$ (see the proof of Lemma 9.8), and $\big|f(x', \pm R)\big| \leqslant g(x')$ for each $x' \in \mathbb{R}^{n-1}$ and $R > 1$.

Therefore, the Dominated Convergence Theorem (Theorem 6.102) implies that

$$\lim_{R\to\infty} \int_{[-R,R]^{n-1}} f(x', \pm R) e^{-i(x',R)\cdot\xi}\, dx' = 0;$$

thus Fubini's Theorem and integrating by parts formula imply that

$$\begin{aligned}
\mathscr{F}\Big[\frac{1}{i}\frac{\partial f}{\partial x_n}(x)\Big](\xi) &= \frac{1}{i}\frac{1}{\sqrt{2\pi}^n} \lim_{R\to\infty} \int_{[-R,R]^n} \frac{\partial f}{\partial x_n}(x) e^{-ix\cdot\xi} dx \\
&= \frac{1}{i}\frac{1}{\sqrt{2\pi}^n} \lim_{R\to\infty} \int_{[-R,R]^{n-1}} \Big(\int_{-R}^{R} \frac{\partial f}{\partial x_n}(x) e^{-ix\cdot\xi} dx_n\Big) dx' \\
&= \frac{1}{i}\frac{1}{\sqrt{2\pi}^n} \lim_{R\to\infty} \Big[\Big(\int_{[-R,R]^{n-1}} f(x', x_n) e^{-i(x',x_n)\cdot\xi}\, dx'\Big)\Big|_{x_n=-R}^{x_n=R} + i\xi_n \int_{[-R,R]^n} f(x) e^{-ix\cdot\xi} dx\Big] \\
&= \xi_n \frac{1}{\sqrt{2\pi}^n} \lim_{R\to\infty} \int_{[-R,R]^n} f(x) e^{-ix\cdot\xi} dx = \xi_k \widehat{f}(\xi).
\end{aligned}$$

□

Corollary 9.12. $\mathcal{P}(\xi_1, \cdots, \xi_n)\widehat{f}(\xi) = \mathscr{F}_x\Big[\mathcal{P}\Big(\frac{1}{i}\frac{\partial}{\partial x_1}, \cdots, \frac{1}{i}\frac{\partial}{\partial x_n}\Big) f(x)\Big](\xi)$ *for all* $f \in \mathscr{S}(\mathbb{R}^n)$ *and polynomial* $\mathcal{P}$.

Corollary 9.13. *The Fourier transform of a Schwartz function is a Schwartz function; that is,* $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$.

Proof. Let $\mathcal{P}$ be a polynomial and $\alpha = (\alpha_1, \cdots, \alpha_n)$ be a multi-index. By Corollary 9.10 and 9.12,

$$\begin{aligned}
\mathcal{P}(\xi) D^\alpha \widehat{f}(\xi) &\equiv \mathcal{P}(\xi_1, \cdots, \xi_n) \frac{\partial^{|\alpha|} \widehat{f}}{\partial \xi_1^{\alpha_1} \cdots \partial \xi_n^{\alpha_n}}(\xi) \\
&= \frac{1}{i^{|\alpha|}} \mathscr{F}_x\Big[\mathcal{P}\big(\frac{1}{i}\frac{\partial}{\partial x_1}, \cdots, \frac{1}{i}\frac{\partial}{\partial x_n}\big)\big[x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} f(x)\big]\Big](\xi);
\end{aligned}$$

thus $\mathcal{P}D^\alpha \widehat{f}$ is the Fourier transform of a Schwartz function g defined by

$$g(x) = \frac{1}{i^{|\alpha|}}\mathcal{P}\Big(\frac{1}{i}\frac{\partial}{\partial x_1}, \cdots, \frac{1}{i}\frac{\partial}{\partial x_n}\Big)\big[x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n} f(x)\big].$$

By Proposition 9.4 and Lemma 9.8, $\mathcal{P}D^\alpha \widehat{f}$ is bounded. □

Remark 9.14. There exists a duality under $\wedge$ between differentiability and rapid decrease: the more differentiability f possesses, the more rapid decrease $\widehat{f}$ has and vice versa.

Definition 9.15. For all $f \in L^1(\mathbb{R}^n)$, we define operator $\mathscr{F}^*$ by

$$(\mathscr{F}^* f)(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi}d\xi.$$

The function $\mathscr{F}^* f$ sometimes is also denoted by $\check{f}$.

The operator $\mathscr{F}^*$, indicated implicitly by the way it is written, is the formal adjoint of $\mathscr{F}$. To be more precise, we have the following

Lemma 9.16. $\langle \mathscr{F}u, v\rangle_{L^2(\mathbb{R}^n)} = \langle u, \mathscr{F}^* v\rangle_{L^2(\mathbb{R}^n)}$ *for all* $u, v \in \mathscr{S}(\mathbb{R}^n)$*, where* $\langle\cdot,\cdot\rangle_{L^2(\mathbb{R}^n)}$ *is an inner product on* $\mathscr{S}(\mathbb{R}^n)$ *given by*

$$\langle u, v\rangle_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n} u(x)\overline{v(x)}\,dx.$$

Proof. Let $u, v \in \mathscr{S}(\mathbb{R}^n)$ be given, and define $f(x,y) = u(x)\overline{v(y)}e^{-ix\cdot y}$. By Tonelli's Theorem (Theorem 6.106), f is absolutely integrable on $\mathbb{R}^n \times \mathbb{R}^n$; thus Fubini's Theorem (Theorem 6.107) implies that

$$\begin{aligned}
\langle \mathscr{F}u, v\rangle_{L^2(\mathbb{R}^n)} &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} u(x)e^{-ix\cdot\xi}dx\Big)\overline{v(\xi)}\,d\xi \\
&= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} u(x)\overline{e^{ix\cdot\xi}v(\xi)}\,d\xi\,dx \\
&= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} u(x)\int_{\mathbb{R}^n}\overline{e^{ix\cdot\xi}v(\xi)}\,d\xi\,dx = \langle u, \mathscr{F}^* v\rangle_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

□

9.3 The Fourier Inversion Formula

We remind the readers that our goal is to prove (9.0.2), while having introduced operators $\mathscr{F}$ and $\mathscr{F}^*$, it is the same as showing that $\mathscr{F}$ and $\mathscr{F}^*$ are inverse to each other; that is, we want to show that

$$\mathscr{F}\mathscr{F}^* = \mathscr{F}^*\mathscr{F} = \mathrm{Id} \qquad \text{on} \quad \mathscr{S}(\mathbb{R}^n).$$

For a given $t > 0$, let $P_t : \mathbb{R} \to \mathbb{R}$ be defined by $P_t(x) = \dfrac{1}{\sqrt{t}} \exp\big(-\dfrac{x^2}{2t}\big)$. Note that $P_t \in \mathscr{S}(\mathbb{R})$ for all $t > 0$, and P_t is normalized so that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} P_t(x)\, dx = 1\,.$$

Now we compute the Fourier transform of P_t. By Lemma 9.9, we find that

$$\begin{aligned}\frac{d\widehat{P}_t}{d\xi}(\xi) &= \frac{-i}{\sqrt{2\pi t}} \int_{\mathbb{R}} xP_t(x) e^{-ix\xi}\, dx \\ &= \frac{-i}{\sqrt{2\pi t}} \int_{\mathbb{R}} xP_t(x) \cos(\xi x)\, dx - \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} xP_t(x) \sin(\xi x)\, dx\,.\end{aligned}$$

Since the functions $y = xP_t(x)$ is absolutely integrable on $\mathbb{R}$ for each fixed $t > 0$, the integral $\displaystyle\int_{\mathbb{R}} xP_t(x) \cos(\xi x)\, dx$ converges absolutely; thus by the fact that $x\cos(\xi x)$ are odd functions in x, we have

$$\int_{\mathbb{R}} xP_t(x) \cos(\xi x)\, dx = \lim_{R\to\infty} \int_{-R}^{R} xP_t(x) \cos(\xi x)\, dx = 0\,.$$

As a consequence,

$$\frac{d\widehat{P}_t}{d\xi}(\xi) = -\frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} xe^{-\frac{x^2}{2t}} \sin(x\xi) dx\,.$$

Integrating by parts,

$$\begin{aligned}\frac{d\widehat{P}_t}{d\xi}(\xi) &= -\frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} xe^{-\frac{x^2}{2t}} \sin(x\xi) dx = -\frac{1}{\sqrt{2\pi t}} \lim_{R\to\infty} \int_{-R}^{R} xe^{-\frac{x^2}{2t}} \sin(x\xi) dx \\ &= -\frac{1}{\sqrt{2\pi t}} \lim_{R\to\infty} \Big[-te^{-\frac{x^2}{2t}} \sin(x\xi)\Big|_{x=-R}^{x=R} + \int_{-R}^{R} \xi te^{-\frac{x^2}{2t}} \cos(x\xi)\, dx \Big] \\ &= -\frac{\xi t}{\sqrt{2\pi t}} \lim_{R\to\infty} \int_{-R}^{R} e^{-\frac{x^2}{2t}} \cos(x\xi)\, dx = -\frac{\xi t}{\sqrt{2\pi t}} \lim_{R\to\infty} \int_{-R}^{R} e^{-\frac{x^2}{2t}} \big[\cos(x\xi) - i\sin(x\xi)\big]\, dx \\ &= -\frac{\xi t}{\sqrt{2\pi t}} \int_{\mathbb{R}} e^{-\frac{x^2}{2t}} e^{-ix\xi} dx = -\xi t \widehat{P}_t(\xi)\,,\end{aligned}$$

thus $\widehat{P}_t(\xi) = Ce^{-\frac{t\xi^2}{2}}$. By the fact that $\widehat{P}_t(0) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} P_t(x) dx = 1$, we must have

$$\widehat{P}_t(\xi) = e^{-\frac{1}{2}t\xi^2}\,. \tag{9.3.1}$$

For $x \in \mathbb{R}^n$, if we define $\mathrm{P}_t(x) = \prod\limits_{k=1}^{n} P_t(x_k) = \Big(\dfrac{1}{\sqrt{t}}\Big)^n e^{-\frac{|x|^2}{2t}}$, then (9.3.1) and Fubini's

Theorem imply that $\widehat{\mathrm{P}}_t(\xi) = e^{-\frac{1}{2}t|\xi|^2}$. Therefore,

$$\widehat{\mathrm{P}}_t(\xi) = \Big(\frac{1}{\sqrt{t}}\Big)^n \mathrm{P}_{\frac{1}{t}}(\xi)$$

which, together with the fact that $\check{f}(x) = \widehat{f}(-x)$, further shows that

$$\check{\widehat{\mathrm{P}}}_t(x) = \Big(\frac{1}{\sqrt{t}}\Big)^n \widehat{\mathrm{P}_{\frac{1}{t}}}(-x) = \Big(\frac{1}{\sqrt{t}}\Big)^n \Big(\frac{1}{\sqrt{t^{-1}}}\Big)^n \mathrm{P}_t(-x) = \mathrm{P}_t(x)\,.$$

Similarly, $\widehat{\check{\mathrm{P}}}_t(\xi) = \mathrm{P}_t(\xi)$, so we establish that

$$\mathscr{F}^*\mathscr{F}(\mathrm{P}_t) = \mathscr{F}\mathscr{F}^*(\mathrm{P}_t) = \mathrm{P}_t\,. \tag{9.3.2}$$

The proof of the following lemma is similar to that of Theorem 8.20.

Lemma 9.17. *If $g \in \mathscr{S}(\mathbb{R}^n)$, then $\mathrm{P}_t \star g \to g$ uniformly on $\mathbb{R}^n$ as $t \to 0^+$, where the convolution operator $\star$ is given by*

$$(\mathrm{P}_t \star g)(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} \mathrm{P}_t(x-y)g(y)\,dy = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} \mathrm{P}_t(y)g(x-y)\,dy\,. \tag{9.3.3}$$

Proof. Let $\varepsilon > 0$ be given. Since $g \in \mathscr{S}(\mathbb{R}^n)$, g is uniformly continuous; thus there exists $\delta > 0$ such that

$$\big|g(x) - g(y)\big| < \frac{\varepsilon}{2} \quad \text{whenever} \quad |x-y| < \delta\,.$$

Since $\displaystyle\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} \mathrm{P}_t(x)\,dx = 1$, for all $x \in \mathbb{R}^n$ we have

$$\begin{aligned}\big|(\mathrm{P}_t \star g)(x) - g(x)\big| &= \frac{1}{\sqrt{2\pi}^n}\Big|\int_{\mathbb{R}^n} g(x-y)\mathrm{P}_t(y)\,dy - \int_{\mathbb{R}^n} g(x)\mathrm{P}_t(y)\,dy\Big| \\ &= \frac{1}{\sqrt{2\pi}^n}\Big|\int_{\mathbb{R}^n} \big[(g(x-y) - g(x)\big]\mathrm{P}_t(y)\,dy\Big| \\ &\leqslant \frac{\varepsilon}{2}\frac{1}{\sqrt{2\pi}^n}\int_{|y|<\delta} \mathrm{P}_t(y)\,dy + \frac{2\|g\|_\infty}{\sqrt{2\pi}^n}\int_{|y|\geqslant\delta} \mathrm{P}_t(y)\,dy\,,\end{aligned}$$

so we obtain that

$$\sup_{x\in\mathbb{R}^n}\big|(\mathrm{P}_t \star g)(x) - g(x)\big| \leqslant \frac{\varepsilon}{2} + \frac{2\|g\|_\infty}{\sqrt{2\pi}^n}\int_{|y|\geqslant\delta} \mathrm{P}_t(y)\,dy\,.$$

Note that

$$\int_{|y|>\delta} \mathrm{P}_t(y)\,dy = \frac{1}{\sqrt{t}^n}\int_{|y|>\delta} e^{-\frac{|y|^2}{2t}}\,dy = \int_{|z|>\frac{\delta}{\sqrt{t}}} e^{-\frac{|z|^2}{2}}\,dz = \int_{\mathbb{R}^n} \mathbf{1}_{B[0,\frac{\delta}{\sqrt{t}}]^\complement}(z)e^{-\frac{|z|^2}{2}}\,dz$$

which, by the Dominated Convergence Theorem, approaches 0 as $t \to 0^+$; thus there exists $h > 0$ such that if $0 < t < h$,

$$\frac{2\|g\|_\infty}{\sqrt{2\pi}^n}\int_{|y|\geqslant\delta} \mathrm{P}_t(y)\,dy < \frac{\varepsilon}{2}.$$

Therefore, we conclude that

$$\sup_{x\in\mathbb{R}^n}\big|(\mathrm{P}_t \star g)(x) - g(x)\big| < \varepsilon \quad \text{whenever} \quad 0 < t < h$$

which shows that $\mathrm{P}_t \star g \to g$ uniformly as $t \to 0^+$. □

Lemma 9.18. *If f and $g \in \mathscr{S}(\mathbb{R}^n)$, then*

$$(\check{f} \star g)(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi}\widehat{g}(\xi)\,d\xi.$$

Proof. By definition of $\check{f}$ and the convolution,

$$(\check{f} \star g)(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\check{f}(x-y)g(y)\,dy = \Big(\frac{1}{2\pi}\Big)^n\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(\xi)e^{i(x-y)\cdot\xi}g(y)\,d\xi\Big)dy.$$

By Tonelli's Theorem (Theorem 6.106), the function $h(\xi, y) = f(\xi)g(y)e^{i(x-y)\cdot\xi}$ is absolutely integrable on $\mathbb{R}^n \times \mathbb{R}^n$; thus Fubini's Theorem (Theorem 6.107) implies that

$$\begin{aligned}(\check{f} \star g)(x) &= \Big(\frac{1}{2\pi}\Big)^n\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi}e^{-iy\cdot\xi}g(y)\,dy\Big)d\xi \\ &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi}\Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} e^{-iy\cdot\xi}g(y)\,dy\Big)d\xi = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi}\widehat{g}(\xi)\,d\xi.\end{aligned}$$ □

Theorem 9.19 (Fourier Inversion Formula)**.** *If $g \in \mathscr{S}(\mathbb{R}^n)$, then $\check{\widehat{g}}(\xi) = \widehat{\check{g}}(\xi) = g(\xi)$. In other words, $\mathscr{F}\mathscr{F}^* = \mathscr{F}^*\mathscr{F} = \mathrm{Id}$.*

Proof. Applying Lemma 9.18 with $f(\xi) = \widehat{\mathrm{P}}_t(\xi) = e^{-\frac{1}{2}t|\xi|^2}$ and using (9.3.2), we find that

$$(\mathrm{P}_t \star g)(x) = (\check{f} \star g)(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} e^{-\frac{1}{2}t|\xi|^2}e^{ix\cdot\xi}\widehat{g}(\xi)\,d\xi.$$

Passing to the limit as $t \to 0^+$, by Lemma 9.17 and Dominated Convergence Theorem (Theorem 6.102) we obtain that

$$g(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\widehat{g}(\xi)e^{ix\cdot\xi}d\xi = \check{\widehat{g}}(x).$$

Let $\widetilde{\ }$ denote the reflection operator given by $\widetilde{f}(x) = f(-x)$. Then the change of variable formula implies that

$$\begin{aligned}\check{g}(\xi) &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(x)e^{ix\cdot\xi}dx = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(x)e^{-i(-x)\cdot\xi}dx \\ &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(-x)e^{-ix\cdot\xi}dx = \widehat{\widetilde{g}}(\xi).\end{aligned}$$

On the other hand,

$$\check{g}(\xi) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(x)e^{-ix\cdot(-\xi)}dx = \widehat{g}(-\xi) = \widetilde{\widehat{g}}(\xi);$$

thus $\widehat{\widehat{g}}(\xi) = \widehat{\widehat{\widetilde{\widetilde{g}}}}(\xi) = \check{\widehat{\widetilde{g}}}(\xi) = g(\xi)$. □

Corollary 9.20. $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ *is a bijection.*

Remark 9.21. In view of the Fourier Inversion Formula (Theorem 9.19), $\mathscr{F}^*$ sometimes is written as $\mathscr{F}^{-1}$, and is called the ***inverse Fourier transform***.

Remark 9.22. In most of the engineering applications the Fourier transform of a function f is defined by

$$\mathscr{F}[f](\xi) = \int_{\mathbb{R}^n} f(x)e^{-ix\cdot\xi}\,dx\,.$$

In this case, the corresponding inverse Fourier transform $\mathscr{F}^{-1}$ and the adjoint Fourier transform $\mathscr{F}^*$ are given by

$$\mathscr{F}^{-1}[f](\xi) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(x)e^{ix\cdot\xi}\,dx \qquad \text{and} \qquad \mathscr{F}^*[f](\xi) = \int_{\mathbb{R}^n} f(x)e^{ix\cdot\xi}\,dx$$

so that $\mathscr{F}^{-1} \neq \mathscr{F}^*$. In some applied fields such as the signal processing the Fourier transform of a function f is defined by

$$\mathscr{F}[f](\xi) = \int_{\mathbb{R}^n} f(x)e^{-2\pi ix\cdot\xi}\,dx\,.$$

In such a case, the inverse Fourier transform and the adjoint Fourier transform are identical and are given by

$$\mathscr{F}^{-1}[f](\xi) = \mathscr{F}^*[f](\xi) = \int_{\mathbb{R}^n} f(x)e^{2\pi ix\cdot\xi}\,dx\,.$$

The proof of this fact is left as an exercise.

Theorem 9.23 (Plancherel formula for $\mathscr{S}(\mathbb{R}^n)$)**.** *If $f,\ g \in \mathscr{S}(\mathbb{R}^n)$, then*

$$\langle f, g\rangle_{L^2(\mathbb{R}^n)} = \langle \widehat{f}, \widehat{g}\rangle_{L^2(\mathbb{R}^n)}\,.$$

Proof. Recall that $\langle f, g\rangle_{L^2(\mathbb{R}^n)} = \displaystyle\int_{\mathbb{R}^n} f(x)\overline{g(x)}\,dx$. By Fubini's theorem,

$$\begin{aligned}\langle \check{f}, g\rangle_{L^2(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \check{f}(x)\overline{g(x)}\,dx = \int_{\mathbb{R}^n}\Big[\frac{1}{\sqrt{2\pi}^{\,n}}\int_{\mathbb{R}^n} f(\xi)e^{ix\cdot\xi}d\xi\Big]\overline{g(x)}\,dx\\ &= \int_{\mathbb{R}^n} f(\xi)\Big[\frac{1}{\sqrt{2\pi}^{\,n}}\int_{\mathbb{R}^n}\overline{g(x)e^{-ix\cdot\xi}dx}\Big]d\xi = \langle f, \widehat{g}\rangle_{L^2(\mathbb{R}^n)}\,.\end{aligned}$$

Therefore, $\langle f, g\rangle_{L^2(\mathbb{R}^n)} = \langle \check{\widehat{f}}, g\rangle_{L^2(\mathbb{R}^n)} = \langle \widehat{f}, \widehat{g}\rangle_{L^2(\mathbb{R}^n)}$. □

Remark 9.24. The Plancherel formula is a "generalization" of the Parseval identity in the following sense. Define the ℓ^2 space as the collection of all square summable (complex) sequences; that is,

$$\ell^2 = \Big\{\{a_k\}_{k=-\infty}^{\infty} \subseteq \mathbb{C}\,\Big|\, \sum_{k=-\infty}^{\infty}|a_k|^2 < \infty\Big\}$$

with inner product

$$\Big\langle \{a_k\}_{k=-\infty}^{\infty}, \{b_k\}_{k=-\infty}^{\infty}\Big\rangle_{\ell^2} = \sum_{k=-\infty}^{\infty} a_k\overline{b_k}\,.$$

Define $\mathcal{F} : L^2(\mathbb{T}) \to \ell^2$ by $\mathcal{F}(f) = \{\widehat{f}_k\}_{k=-\infty}^{\infty}$. Then (8.5.5) shows that

$$\langle f, g\rangle_{L^2(\mathbb{T})} = \langle \mathcal{F}(f), \mathcal{F}(g)\rangle_{\ell^2} \qquad \forall\, f, g \in L^2(\mathbb{T})$$

so that we obtain an identity similar to the Plancherel formula.

Remark 9.25. Even though in general an square integrable function might not be integrable, using the Plancherel formula the Fourier transform of L^2-functions can still be defined. Note that the Plancherel formula provides that

$$\|f\|_{L^2(\mathbb{R}^n)} = \|\widehat{f}\|_{L^2(\mathbb{R}^n)} \qquad \forall\, f \in \mathscr{S}(\mathbb{R}^n)\,. \tag{9.3.4}$$

If $f \in L^2(\mathbb{R}^n)$; that is, $|f|$ is square integrable, by the fact that $\mathscr{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, there exists a sequence $\{f_k\}_{k=1}^{\infty} \subseteq \mathscr{S}(\mathbb{R}^n)$ such that $\lim\limits_{k\to\infty}\|f_k - f\|_{L^2(\mathbb{R}^n)} = 0$. Then $\{f_k\}_{k=1}^{\infty}$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$; thus (9.3.4) implies that $\{\widehat{f}_k\}_{k=1}^{\infty}$ is also a Cauchy sequence in $L^2(\mathbb{R}^n)$. By the completeness of $L^2(\mathbb{R}^n)$ (which we did not cover in this lecture), there exists $g \in L^2(\mathbb{R}^n)$ such that

$$\lim_{k\to\infty} \|\widehat{f}_k - g\|_{L^2(\mathbb{R}^n)} = 0\,.$$

We note that such a limit g is independent of the choice of sequence $\{f_k\}_{k=1}^\infty$ used to approximate f; thus we can denote this limit g as $\widehat{f}$. In other words, $\mathscr{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. Moreover, by that $f_k \to f$ and $\widehat{f}_k \to \widehat{f}$ in $L^2(\mathbb{R}^n)$ as $k \to \infty$, we find that

$$\|f\|_{L^2(\mathbb{R}^n)} = \|\widehat{f}\|_{L^2(\mathbb{R}^n)} \qquad \forall\, f \in L^2(\mathbb{R}^n)\,,$$

and the parallelogram law further implies that $\langle f, g\rangle_{L^2(\mathbb{R}^n)} = \langle \widehat{f}, \widehat{g}\rangle_{L^2(\mathbb{R}^n)}$ for all $f, g \in L^2(\mathbb{R}^n)$. Similar argument applies to the case of inverse transform of L^2-functions; thus we conclude that

$$\langle f, g\rangle_{L^2(\mathbb{R}^n)} = \langle \widehat{f}, \widehat{g}\rangle_{L^2(\mathbb{R}^n)} = \langle \check{f}, \check{g}\rangle_{L^2(\mathbb{R}^n)} \qquad \forall\, f, g \in L^2(\mathbb{R}^n)\,. \tag{9.3.5}$$

We will talk about how to define the Fourier transform of L^2-functions in another way in Section 9.4.

Theorem 9.26. *If $f, g \in \mathscr{S}(\mathbb{R}^n)$, then $\mathscr{F}(f \star g) = \widehat{f}\,\widehat{g}$. In particular, $f \star g \in \mathscr{S}(\mathbb{R}^n)$ if $f, g \in \mathscr{S}(\mathbb{R}^n)$.*

Proof. By the definition of the Fourier transform and the convolution,

$$\begin{aligned}
\widehat{f \star g}(\xi) &= \frac{1}{\sqrt{2\pi}^n}\mathscr{F}\Big(\int_{\mathbb{R}^n} f(\cdot - y)g(y)\,dy\Big)(\xi) \\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\Big[\int_{\mathbb{R}^n} f(x-y)g(y)\,dy\Big]e^{-ix\cdot\xi}dx \\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} g(y)\Big(\int_{\mathbb{R}^n} f(x)e^{-i(x+y)\cdot\xi}dx\Big)dy \\
&= \Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(x)e^{-ix\cdot\xi}dx\Big)\Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(y)e^{-iy\cdot\xi}dy\Big)
\end{aligned}$$

which concludes the theorem. □

Corollary 9.27. *$\mathscr{F}^*(f \star g) = \check{f}\,\check{g}$, $\mathscr{F}(fg) = \widehat{f} \star \widehat{g}$ and $\mathscr{F}^*(fg) = \check{f} \star \check{g}$ for all $f, g \in \mathscr{S}(\mathbb{R}^n)$.*

We have established the Fourier inversion formula for Schwartz class functions. Our goal next is to show that the Fourier inversion formula holds for a larger class of functions. Motivated by the Fourier inversion formula, we would like to show, if possible, that

$$\check{\widehat{f}} = \widehat{\check{f}} = f \qquad \forall\, f \in L^1(\mathbb{R}^n) \text{ such that } \widehat{f} \in L^1(\mathbb{R}^n).$$

The above assertion cannot be true since $\hat{\check{f}}$ and $\check{\hat{f}}$ are both continuous (by Proposition 9.4) while $f \in L^1(\mathbb{R}^n)$ which is not necessary continuous. However, we will prove that the identity above holds at points of continuity of f.

Before proceeding, we establish a lemma which is very similar to Lemma 9.16.

Lemma 9.28. *Let $f \in L^1(\mathbb{R}^n)$ and $g \in \mathscr{S}(\mathbb{R}^n)$. Then $\langle \hat{f}, g\rangle = \langle f, \hat{g}\rangle$ and $\langle \check{f}, g\rangle = \langle f, \check{g}\rangle$, where $\langle f, g\rangle = \int_{\mathbb{R}^n} f(x)g(x)\,dx$.*

Proof. We only prove $\langle \hat{f}, g\rangle = \langle f, \hat{g}\rangle$ if $f \in L^1(\mathbb{R}^n)$ and $g \in \mathscr{S}(\mathbb{R}^n)$. By Proposition 9.4, $\hat{f}$ is bounded and continuous on $\mathbb{R}^n$; thus $\hat{f}g$ is an absolutely integrable continuous function. By Fubini's Theorem,

$$\begin{aligned}\langle \hat{f}, g\rangle &= \int_{\mathbb{R}^n}\Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(x)e^{-ix\cdot\xi}dx\Big)g(\xi)d\xi = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(x)g(\xi)e^{-ix\cdot\xi}dx\Big)d\xi\\ &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} f(x)g(\xi)e^{-ix\cdot\xi}d\xi\Big)dx = \int_{\mathbb{R}^n} f(x)\Big(\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(\xi)e^{-ix\cdot\xi}d\xi\Big)dx\end{aligned}$$

which is exactly $\langle f, \hat{g}\rangle$. ▫

Recall that our goal is to show that

$$\text{if } f, \hat{f} \in L^1(\mathbb{R}^n), \text{ then } \check{\hat{f}}(x) = \hat{\check{f}}(x) = f(x) \text{ whenever } f \text{ is continuous at } x.$$

This amounts to treat $\check{\hat{f}}$ and f in the same vector space and check if $\check{\hat{f}} - f$ is the zero vector in that vector space. This underlying vector space is introduced in the following

Definition 9.29. The space $L^1_{\text{loc}}(\mathbb{R}^n)$ consists of all functions (defined on $\mathbb{R}^n$) that are absolutely integrable on all bounded open balls of $\mathbb{R}^n$. In other words,

$$L^1_{\text{loc}}(\mathbb{R}^n) = \Big\{f : \mathbb{R}^n \to \mathbb{C} \,\Big|\, \int_{B(a,r)} f(x)\,dx \text{ is absolutely convergent for all } a \in \mathbb{R}^n \text{ and } r > 0\Big\}.$$

Again, we emphasize that we ***abuse*** the notation $L^1_{\text{loc}}(\mathbb{R}^n)$ which in fact stands for a larger class of functions. We also note that $L^1(\mathbb{R}^n) \subseteq L^1_{\text{loc}}(\mathbb{R}^n)$ and $\mathscr{C}_b(\mathbb{R}^n;\mathbb{C}) \subseteq L^1_{\text{loc}}(\mathbb{R}^n)$.

How do we determine if an locally integrable function h is the zero vector in $L^1_{\text{loc}}(\mathbb{R}^n)$? Our goal is to establish an equivalent condition of that $h \equiv 0$ (or to be more precisely, $h(x) = 0$ if h is continuous at x) stated briefly as follows (the precise statement is given in Lemma 9.35):

$$\text{If } h \in L^1_{\text{loc}}(\mathbb{R}^n), \text{ then } h \equiv 0 \text{ if and only if } \langle h, g\rangle = 0 \text{ for all } g \in \mathscr{S}(\mathbb{R}^n). \tag{9.3.6}$$

We note that the difficulty here is that h and g belongs to different space so that we cannot simply let $g = \overline{h}$ to conclude that $\displaystyle\int_{\mathbb{R}^n} |h(x)|^2\,dx = 0$.

A special class of functions that will be used as the role of g in (9.3.6) is called the standard mollifiers. Let $\zeta : \mathbb{R} \to \mathbb{R}$ be a smooth function defined by

$$\zeta(x) = \begin{cases} \exp\left(\dfrac{1}{x^2-1}\right) & \text{if } |x| < 1\,, \\ 0 & \text{if } |x| \geqslant 1\,. \end{cases}$$

For $x \in \mathbb{R}^n$, define $\eta_1(x) = C\zeta(|x|)$, where C is chosen so that $\displaystyle\int_{\mathbb{R}^n} \eta_1(x)\,dx = 1$. The change of variables formula then implies that $\eta_\varepsilon(x) \equiv \varepsilon^{-n}\eta_1(x/\varepsilon)$ has integral 1. We remark that η_ε is smooth and $\eta_\varepsilon(x) \neq 0$ if and only if $x \in B(0,\varepsilon)$; thus $\eta_\varepsilon \in \mathscr{S}(\mathbb{R}^n)$ for all $\varepsilon > 0$.

Definition 9.30. The collection of functions $\{\eta_\varepsilon\}_{\varepsilon>0}$ is called the ***standard mollifiers***.

Definition 9.31. Let f, g be functions defined on $\mathbb{R}^n$. The convolution of f and g, denoted by $f * g$, is a function defined on $\mathbb{R}^n$ given by

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy$$

whenever the integral makes sense for all $x \in \mathbb{R}^n$.

We note that the change of variables formula implies that $f * g = g * f$ whenever the convolution makes sense.

Remark 9.32. Let $\{\eta_\varepsilon\}_{\varepsilon>0}$ be the standard mollifiers.

1. Since $\eta_\varepsilon(y) \neq 0$ if and only if $y \in B(0,\varepsilon)$, for each $x \in \mathbb{R}^n$,

$$(\eta_\varepsilon * f)(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)f(y)\,dy = \int_{B(x,\varepsilon)} \eta_\varepsilon(x-y)f(y)\,dy$$

   and the integral exists if $f \in L^1_{\text{loc}}(\mathbb{R}^n)$.

2. By the fact that η_ε is smooth, the Dominated Convergence Theorem (Theorem 6.102) shows that $\eta_\varepsilon * f$ is smooth for each $\varepsilon > 0$, and

$$D^\alpha(\eta_\varepsilon * f)(x) = \int_{\mathbb{R}^n} (D^\alpha\eta)(x-y)f(y)\,dy\,,$$

   where $\alpha = (\alpha_1, \cdots, \alpha_n)$ is a multi-index. The detail proof is left as an exercise.

Example 9.33. Let $f = \mathbf{1}_{[a,b]}$, the characteristic/indicator function of the closed interval $[a,b]$. Then for $\varepsilon \ll 1$, the function $\eta_\varepsilon * f$ is smooth and has the property that

$$(\eta_\varepsilon * f)(x) = \begin{cases} 1 & \text{if } x \in [a+\varepsilon, b-\varepsilon], \\ 0 & \text{if } x \in [a-\varepsilon, b+\varepsilon]^\complement, \end{cases}$$

and $0 \leqslant f \leqslant 1$. Moreover, $\eta_\varepsilon * f$ converges pointwise to f on $\mathbb{R}\backslash\{a,b\}$.

The following lemma shows that $\eta_\varepsilon * f$ converges to f at points of continuity of f if $f \in L^1_{\text{loc}}(\mathbb{R}^n)$.

Lemma 9.34. *Let $f \in L^1_{loc}(\mathbb{R}^n)$ and x_0 be a continuity of f. Then*

$$\lim_{\varepsilon \to 0^+} (\eta_\varepsilon * f)(x_0) = f(x_0).$$

Proof. Let $\epsilon > 0$ be given. Since f is continuous at x_0, there exists $\delta > 0$ such that

$$\big|f(y) - f(x_0)\big| < \frac{\epsilon}{2} \quad \text{whenever} \quad |y - x_0| < \delta.$$

Therefore, by the fact that $\int_{\mathbb{R}^n} \eta_\varepsilon(y)\,dy = 1$, for $0 < \varepsilon < \delta$ we find that

$$\begin{aligned} \big|(\eta_\varepsilon * f)(x_0) - f(x_0)\big| &= \Big|\int_{\mathbb{R}^n} \eta_\varepsilon(y) f(x_0 - y)\,dy - \int_{\mathbb{R}^n} \eta_\varepsilon(y) f(x_0)\,dy\Big| \\ &\leqslant \int_{B(0,\varepsilon)} \eta_\varepsilon(y)\big|f(x_0 - y) - f(x_0)\big|\,dy \leqslant \frac{\epsilon}{2}\int_{B(0,\varepsilon)} \eta_\varepsilon(y)\,dy < \epsilon. \end{aligned}$$ □

Lemma 9.35. *Let $f \in L^1_{\text{loc}}(\mathbb{R}^n)$. If $\langle f, g\rangle = 0$ for all $g \in \mathscr{S}(\mathbb{R}^n)$, then $f(x_0) = 0$ whenever f is continuous at x_0.*

Proof. Let $\{\eta_\varepsilon\}_{\varepsilon>0}$ be the standard mollifiers, and x_0 be a point of continuity of f. Then for all $\varepsilon > 0$,

$$(\eta_\varepsilon * f)(x_0) = \int_{\mathbb{R}^n} \eta_\varepsilon(x_0 - y) f(y)\,dy = \int_{\mathbb{R}^n} f(y)\eta_\varepsilon(y - x_0)\,dy = \langle f, \tau_{x_0}\eta_\varepsilon\rangle,$$

where τ_{x_0} is the translation operator defined by $(\tau_{x_0}\phi)(y) = \phi(y - x_0)$. Since $\eta_\varepsilon \in \mathscr{S}(\mathbb{R}^n)$ for all $\varepsilon > 0$, the function $\tau_{x_0}\eta_\varepsilon \in \mathscr{S}(\mathbb{R}^n)$ for all $\varepsilon > 0$; thus $(\eta_\varepsilon * f)(x_0) = 0$ for all $\varepsilon > 0$. By Lemma 9.34, we conclude that

$$f(x_0) = \lim_{\varepsilon \to 0}(\eta_\varepsilon * f)(x) = 0.$$ □

Remark 9.36. The reverse statement of Lemma 9.35 is also true: if $f \in L^1_{\rm loc}(\mathbb{R}^n)$ has the property that $f(x_0) = 0$ whenever f is continuous at x_0, then $\langle f, g\rangle = 0$ for all $g \in \mathscr{S}(\mathbb{R}^n)$ since if $f \in L^1_{\rm loc}(\mathbb{R}^n)$, the collection of discontinuities of f has measure zero which shows that $f(x) \neq 0$ only on a set of measure zero. Therefore, $\langle f, g\rangle = 0$ for all $g \in \mathscr{S}(\mathbb{R}^n)$. In other words, if $f \in L^1_{\rm loc}(\mathbb{R}^n)$, then

$$\langle f, g\rangle = 0 \text{ for all } g \in \mathscr{S}(\mathbb{R}^n) \text{ if and only if } f(x_0) = 0 \text{ whenever } f \text{ is continuous at } x_0.$$

Lemma 9.35 establishes the non-trival direction "$\Rightarrow$".

Now we are in position of showing the Fourier inversion formula for functions of more general class.

Theorem 9.37 (Fourier Inversion Formula)**.** *Let $f \in L^1(\mathbb{R}^n)$ such that $\widehat{f} \in L^1(\mathbb{R}^n)$. Then*

$$\check{\widehat{f}}(x) = \widehat{\check{f}}(x) = f(x) \qquad \textit{whenever } f \textit{ is continuous at } x.$$

Proof. Let $f : \mathbb{R}^n \to \mathbb{C}$ be such that $f, \widehat{f} \in L^1(\mathbb{R}^n)$. By the fact that $\check{f}(\xi) = \widehat{f}(-\xi)$ for all $\xi \in \mathbb{R}^n$, the change of variables formula implies that $\check{f} \in L^1(\mathbb{R}^n)$.

Let $g \in \mathscr{S}(\mathbb{R}^n)$ be given. By Lemma 9.28 and the Fourier inversion formula for Schwartz class functions (Theorem 9.19),

$$\langle \check{\widehat{f}}, g\rangle = \langle \widehat{f}, \check{g}\rangle = \langle f, \widehat{\check{g}}\rangle = \langle f, g\rangle \qquad \text{and} \qquad \langle \widehat{\check{f}}, g\rangle = \langle \check{f}, \widehat{g}\rangle = \langle f, \check{\widehat{g}}\rangle = \langle f, g\rangle.$$

In other words, if $f, \widehat{f} \in L^1(\mathbb{R}^n)$,

$$\langle \check{\widehat{f}} - f, g\rangle = \langle \widehat{\check{f}} - f, g\rangle = 0 \qquad \forall\, g \in \mathscr{S}(\mathbb{R}^n).$$

By Proposition 9.4, $\check{\widehat{f}}, \widehat{\check{f}} \in \mathscr{C}_b(\mathbb{R}^n; \mathbb{C})$; thus the collection of points of continuities of $\check{\widehat{f}} - f$ and $\widehat{\check{f}} - f$ is identical to the collection of points of continuities of f; that is,

$$f \text{ is continuous at } x \text{ if and only if } \check{\widehat{f}} - f \text{ and } \widehat{\check{f}} - f \text{ are continuous at } x.$$

Moreover, by the fact that $\mathscr{C}_b(\mathbb{R}^n; \mathbb{C}) \subseteq L^1_{\rm loc}(\mathbb{R}^n)$ and $L^1(\mathbb{R}^n) \subseteq L^1_{\rm loc}(\mathbb{R}^n)$ we find that $\check{\widehat{f}} - f, \widehat{\check{f}} - f \in L^1_{\rm loc}(\mathbb{R}^n)$. Therefore, the theorem is concluded by Lemma 9.35. □

Remark 9.38. Since an integrable function $f : \mathbb{R}^n \to \mathbb{R}$ must be continuous **almost everywhere** on $\mathbb{R}^n$, Theorem 9.37 implies that if $f : \mathbb{R}^n \to \mathbb{R}$ is a function such that f, $\widehat{f} \in L^1(\mathbb{R}^n)$, then $\check{\widehat{f}} = \widehat{\check{f}} = f$ almost everywhere.

9.4 The Fourier Transform of Generalized Functions

It is often required to consider the Fourier transform of functions which do not belong to $L^1(\mathbb{R}^n)$. For example, the **normalized sinc function** $\mathrm{sinc} : \mathbb{R} \to \mathbb{R}$ defined by

$$\mathrm{sinc}(x) = \begin{cases} \dfrac{\sin(\pi x)}{\pi x} & \text{if } x \neq 0\,, \\ 1 & \text{if } x = 0\,, \end{cases} \tag{9.4.1}$$

does not belong to $L^1(\mathbb{R})$ but it is a very important function in the study of signal processing.

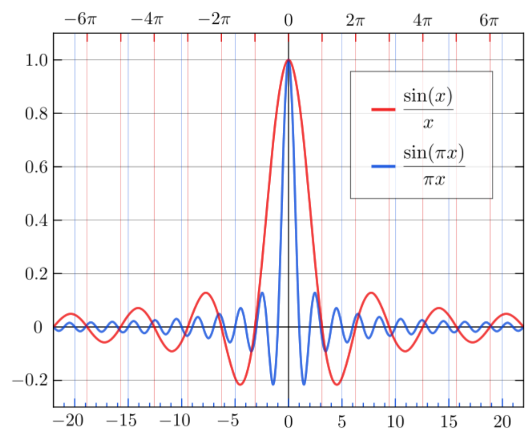


Figure 9.1: The graphs of unnormalized and normalized sinc functions (from wiki)

Moreover, there are "functions" that are not even functions in the traditional sense. For example, in physics and engineering applications the Dirac delta "function" δ is defined as the "function" which validates the relation

$$\int_{\mathbb{R}^n} \delta(x)\phi(x)\,dx = \phi(0) \qquad \forall\, \phi \in \mathscr{C}(\mathbb{R}^n)$$

In fact, there is no function (in the traditional sense) satisfying the property given above (reasoning later). Can we take the Fourier transform of those "functions" as well? To understand this topic better, it is required to study the theory of generalized functions/ distributions.

To understand the meaning of distributions, let us turn to a situation in physics: measuring the temperature. To measure the temperature T at a point x, instead of outputting the exact value of $T(x)$ the thermometer instead outputs the **overall value** of the temperature near x. In other words, the reading of the temperature is determined by a pairing of the temperature distribution with the thermometer.

In mathematical point of views, to evaluate the function value of a locally integrable function f at a point of continuity x_0, we apply Lemma 9.34 and obtain that

$$f(x_0) = \lim_{\varepsilon \to 0^+} (\eta_\varepsilon * f)(x_0) = \lim_{\varepsilon \to 0^+} \big\langle f, \tau_{x_0}\eta_\varepsilon \big\rangle, \tag{9.4.2}$$

where τ_{x_0} is the translation operator given by $(\tau_{x_0}\phi)(x) = \phi(x-x_0)$. Here η_ε can be viewed as a meter that can measure the function value of locally integrable functions, ε is a parameter that corresponds to the accuracy of this meter, and $\tau_{x_0}\eta_\varepsilon$ is a meter that locates at position x_0. Nevertheless, for $f \in L^1_{\rm loc}(\mathbb{R}^n)$, the "pairing" $\langle f, \phi \rangle$ is defined not only on functions of the form $\phi = \tau_{x_0}\eta_\varepsilon$ but also on $\phi \in \mathscr{D}(\mathbb{R}^n)$, where

$$\mathscr{D}(\mathbb{R}^n) \equiv \Big\{\phi : \mathbb{R}^n \to \mathbb{C} \,\Big|\, \phi \in \mathscr{C}^\infty(\mathbb{R}^n) \text{ and } \operatorname{supp}(\phi) \equiv \overline{\{x \in \mathbb{R}^n \; \phi(x) \neq 0\}} \text{ is compact}\Big\}.$$

This pairing induced a (continuous) linear functional $T_f : \mathscr{D}(\mathbb{R}^n) \to \mathbb{C}$ defined by

$$T_f(\phi) = \langle f, \phi \rangle. \tag{9.4.3}$$

Moreover, if $T : \mathscr{D}(\mathbb{R}^n) \to \mathbb{C}$ is a continuous linear functional, and there exists $f \in L^1_{\rm loc}(\mathbb{R}^n)$ such that $T(\phi) = \langle f, \phi \rangle$ for all $\phi \in \mathscr{D}(\mathbb{R}^n)$, then f is uniquely determined except perhaps on a set of measure zero (or to be more precise, f is uniquely determined at all points of continuity of f). Therefore, with $\mathscr{D}(\mathbb{R}^n)'$ denoting the collection of all (continuous) linear functionals defined on $\mathscr{D}(\mathbb{R}^n)$, there is a natural injection $\iota : L^1_{\rm loc}(\mathbb{R}^n) \to \mathscr{D}(\mathbb{R}^n)'$ (given by $\iota(f) = T_f$).

On the other hand, a (continuous) linear functional defined on $\mathscr{D}(\mathbb{R}^n)$ might not take the form of (9.4.3). For example, there exists **no** locally integrable function f such that

$$T(\phi) = \int_{\mathbb{R}^n} f(y)\phi(y)\, dy = \phi(0) \qquad \forall\, \phi \in \mathscr{D}(\mathbb{R}^n). \tag{9.4.4}$$

To see this, suppose the contrary that there exists $f \in L^1_{\rm loc}(\mathbb{R}^n)$ such that (9.4.4) holds. Then

$$(\eta_\varepsilon * f)(x) = \int_{\mathbb{R}^n} f(y)(\tau_x\eta_\varepsilon)(y)\, dy = \eta_\varepsilon(x)$$

which vanishes if $x \notin B(0,\varepsilon)$. This implies that $f = 0$ almost everywhere so that $\langle f, \phi \rangle = 0$ for all $\phi \in \mathscr{D}(\mathbb{R}^n)$, a contradiction. Therefore, $\iota : L^1_{\rm loc}(\mathbb{R}^n) \to \mathscr{D}(\mathbb{R}^n)'$ (given by $\iota(f) = T_f$) is not surjective; thus (continuous) linear functionals on $\mathscr{D}(\mathbb{R}^n)$ defines more "functions" than $L^1_{\rm loc}(\mathbb{R}^n)$. Such kind of linear functionals are called generalized functions or distributions.

The Fourier transform can be defined on a smaller class of generalized functions, the space of tempered distributions. A tempered distribution is a continuous linear functional on $\mathscr{S}(\mathbb{R}^n)$. In other words, T is a tempered distribution if

$$T : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C},\ T(c\phi + \psi) = cT(\phi) + T(\psi) \text{ for all } c \in \mathbb{C} \text{ and } \phi, \psi \in \mathscr{S}(\mathbb{R}^n),$$
$$\text{and } \lim_{j\to\infty} T(\phi_j) = T(\phi) \text{ if } \{\phi_j\}_{j=1}^{\infty} \subseteq \mathscr{S}(\mathbb{R}^n) \text{ and } \phi_j \to \phi \text{ in } \mathscr{S}(\mathbb{R}^n).$$

The convergence in $\mathscr{S}(\mathbb{R}^n)$ is described by semi-norms, and is given in the following

Definition 9.39 (Convergence in $\mathscr{S}(\mathbb{R}^n)$)**.** For each $k \in \mathbb{N} \cup \{0\}$, define the semi-norm

$$p_k(\phi) = \sup_{x\in\mathbb{R}^n, |\alpha|\leqslant k} \langle x\rangle^k |D^\alpha \phi(x)|,$$

where $\langle x\rangle = (1 + |x|^2)^{\frac{1}{2}}$. A sequence $\{\phi_j\}_{j=1}^{\infty} \subseteq \mathscr{S}(\mathbb{R}^n)$ is said to converge to ϕ in $\mathscr{S}(\mathbb{R}^n)$ if $p_k(\phi_j - \phi) \to 0$ as $j \to \infty$ for all $k \in \mathbb{N} \cup \{0\}$.

We note that $p_k(\phi) \leqslant p_{k+1}(\phi)$, so $\{\phi_j\}_{j=1}^{\infty} \subseteq \mathscr{S}(\mathbb{R}^n)$ converges to ϕ in $\mathscr{S}(\mathbb{R}^n)$ whenever $p_k(\phi_j - \phi) \to 0$ as $j \to \infty$ for $k \gg 1$. We also note that if $\{\phi_j\}_{j=1}^{\infty}$ converge to ϕ in $\mathscr{S}(\mathbb{R}^n)$, then $\{\phi_j\}_{j=1}^{\infty}$ converges uniformly to ϕ on $\mathbb{R}^n$.

Definition 9.40 (Tempered Distributions)**.** A linear map $T : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ is continuous if there exists $N \in \mathbb{N}$ such that for each $k \geqslant N$, there exists a constant C_k such that

$$\big|T(\phi)\big| \leqslant C_k p_k(\phi) \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).$$

The collection of continuous linear functionals on $\mathscr{S}(\mathbb{R}^n)$ is denoted by $\mathscr{S}(\mathbb{R}^n)'$. Elements of $\mathscr{S}(\mathbb{R}^n)'$ are called ***tempered distributions***.

Example 9.41. For $1 \leqslant p < \infty$, let $L^p(\mathbb{R}^n)$ denote the collection of Riemann measurable functions whose p-th power is integrable; that is,

$$L^p(\mathbb{R}^n) = \Big\{ f : \mathbb{R}^n \to \mathbb{C} \,\Big|\, f \text{ is Riemann measurable and } \int_{\mathbb{R}^n} \big|f(x)\big|^p\, dx < \infty \Big\},$$

and let $L^\infty(\mathbb{R}^n)$ denote the collection of bounded Riemann measurable functions. Every L^p-function $f : \mathbb{R}^n \to \mathbb{C}$ can be viewed as a tempered distribution for all $p \in [1, \infty]$. In fact, the tempered distribution T_f associated with f is defined by

$$T_f(\phi) = \int_{\mathbb{R}^n} f(x)\phi(x)\, dx \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n). \tag{9.4.5}$$

Now we show that T_f given by (9.4.5) is indeed a tempered distribution. Let $\phi \in \mathscr{S}(\mathbb{R}^n)$ be given. Then $\|\phi\|_{L^\infty(\mathbb{R}^n)} \leqslant p_k(\phi)$ for all $k \in \mathbb{N}$, while for $1 \leqslant q < \infty$ and $k > \dfrac{n}{q}$,

$$\begin{aligned}\|\phi\|_{L^q(\mathbb{R}^n)} &\equiv \Big(\int_{\mathbb{R}^n} \big|\phi(x)\big|^q\,dx\Big)^{\frac{1}{q}} = \Big(\int_{\mathbb{R}^n} \langle x\rangle^{-kq}\big[\langle x\rangle^k|\phi(x)|\big]^q\,dx\Big)^{\frac{1}{q}} \leqslant \Big(\int_{\mathbb{R}^n} \langle x\rangle^{-kq}\,dx\Big)^{\frac{1}{q}} p_k(\phi)\\ &\leqslant \Big(\omega_{n-1}\int_0^\infty (1+r^2)^{-\frac{kq}{2}} r^{n-1}dr\Big)^{\frac{1}{q}} p_k(\phi).\end{aligned}$$

Note that $\displaystyle\int_0^\infty (1+r^2)^{-\frac{kq}{2}} r^{n-1}dr < \infty$ if $k > \dfrac{n}{q}$; thus for all $q \in [1,\infty]$, there exists $C_{k,q,n} > 0$ such that

$$\|\phi\|_{L^q(\mathbb{R}^n)} \leqslant C_{k,q,n}p_k(\phi) \qquad \forall\, k \gg 1. \tag{9.4.6}$$

Therefore, if $f \in L^p(\mathbb{R}^n)$, by the Hölder inequality we have

$$\big|\langle f,\phi\rangle\big| \leqslant \|f\|_{L^p(\mathbb{R}^n)}\|\phi\|_{L^{p'}(\mathbb{R}^n)} \leqslant C_{k,p',n}\|f\|_{L^p(\mathbb{R}^n)}p_k(\phi) \qquad \forall\, k \gg 1,$$

where $p' \in [1,\infty]$ is the Hölder conjugate of p satisfying $\dfrac{1}{p} + \dfrac{1}{p'} = 1$; thus $T_f \in \mathscr{S}(\mathbb{R}^n)'$ if $f \in L^p(\mathbb{R}^n)$. Note that the sinc function belongs to $L^2(\mathbb{R})$ so that $T_{\rm sinc} \in \mathscr{S}(\mathbb{R})'$.

Example 9.42. Let $f : \mathbb{R} \to \mathbb{R}$ be a 2π-periodic, Riemann measurable function such that $\displaystyle\int_{-\pi}^{\pi} |f(x)|\,dx < \infty$, and $\phi \in \mathscr{S}(\mathbb{R})$. By the definition of the p_k semi-norm,

$$|x|^2\big|\phi(x)\big| \leqslant p_2(\phi) \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n) \text{ and } x \in \mathbb{R}^n.$$

Therefore,

$$\begin{aligned}\big|\langle f,\phi\rangle\big| &= \Big|\sum_{k=-\infty}^{\infty}\int_{-\pi+2k\pi}^{\pi+2k\pi} f(x)\phi(x)\,dx\Big| \leqslant \sum_{k=-\infty}^{\infty}\int_{-\pi}^{\pi}\big|f(x)\big|\big|\phi(x+2k\pi)\big|\,dx\\ &= \int_{-\pi}^{\pi}\big|f(x)\big|\big|\phi(x)\big|\,dx + \sum_{|k|\geqslant 1}\int_{-\pi}^{\pi}\big|f(x)\big|\big|\phi(x+2k\pi)\big|\,dx\\ &\leqslant p_0(\phi)\int_{-\pi}^{\pi}\big|f(x)\big|\,dx + \sum_{|k|\geqslant 1}\int_{-\pi}^{\pi}\big|f(x)\big|\frac{1}{|x+2k\pi|^2}p_2(\phi)\,dx\\ &\leqslant \Big(\int_{-\pi}^{\pi}\big|f(x)\big|\,dx\Big)\Big(1 + 2\sum_{k=1}^{\infty}\frac{1}{(2k-1)^2}\Big)p_2(\phi)\end{aligned}$$

which implies that T_f is a tempered distribution. In particular, $T_c \in \mathscr{S}(\mathbb{R})'$ for all constant $c \in \mathbb{R}$.

From now on, we identify f with the tempered distribution T_f if f is an ordinary function (or a function defined pointwise). In other words, if $T \in \mathscr{S}(\mathbb{R}^n)'$ and $f : \mathbb{R}^n \to \mathbb{C}$ is a function, we say that $T = f$ in $\mathscr{S}(\mathbb{R}^n)'$ if $T = T_f$, where T_f is the tempered distribution associated with the function f. Moreover, if $T \in \mathscr{S}(\mathbb{R}^n)'$ and $\phi \in \mathscr{S}(\mathbb{R}^n)$, $T(\phi)$ is also expressed as $\langle T, \phi \rangle$.

Example 9.43 (Dirac delta function)**.** Consider the map $\delta : \mathscr{C}(\mathbb{R}^n) \to \mathbb{R}$ defined by $\delta(\phi) = \phi(0)$. Then $\left|\delta(\phi)\right| \leqslant p_0(\phi) \leqslant p_k(\phi)$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$ and $k \in \mathbb{N} \cup \{0\}$; thus $\delta \in \mathscr{S}(\mathbb{R}^n)'$. Therefore, we also write $\delta(\phi)$ as $\langle \delta, \phi \rangle$. This explains why the Dirac delta function has the property that

$$\int_{\mathbb{R}^n} \delta(x)\phi(x)\,dx = \phi(0)$$

since the integral above is an informal expression of $\langle \delta, \phi \rangle$.

Similarly, the Dirac delta function at a point ω defined by $\langle \delta_\omega, \phi \rangle = \phi(\omega)$ is also a tempered distribution.

Remark 9.44. Not all ordinary functions are tempered distributions. For example, the function $f(x) = e^{x^4}$ is locally integrable (since it is continuous), but $\displaystyle\int_{\mathbb{R}} f(x)e^{-x^2}\,dx = \infty$. Therefore, being in $L^1_{\text{loc}}(\mathbb{R}^n)$ is not good enough to generate elements in $\mathscr{S}(\mathbb{R}^n)'$, and it requires that $|f(x)| \leqslant C(1 + |x|^N)$ for any $N \in \mathbb{N}$. In such a case, $T_f \in \mathscr{S}(\mathbb{R}^n)'$ is well-defined.

As shown in the example above, a tempered distribution might not be defined in the pointwise sense. Therefore, how to define usual operations such as translation, dilation, and reflection on generalized functions should be answered prior to define the Fourier transform of tempered distributions. For completeness, let us start from providing the definitions of translation, dilation and reflection operators.

Definition 9.45 (Translation, dilation, and reflection)**.** Let $f : \mathbb{R}^n \to \mathbb{C}$ be a function.

1. For $h \in \mathbb{R}^n$, the translation operator τ_h maps f to $\tau_h f$ given by $(\tau_h f)(x) = f(x - h)$.

2. For $\lambda > 0$, the dilation operator $d_\lambda : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ maps f to $d_\lambda f$ given by $(d_\lambda f)(x) = f(\lambda^{-1}x)$.

3. The reflection operator $\sim$ maps f to $\widetilde{f}$ given by $\widetilde{f}(x) = f(-x)$.

Now suppose that $T \in \mathscr{S}(\mathbb{R}^n)'$. We expect that $\tau_h T$, $d_\lambda T$ and $\widetilde{T}$ are also tempered distributions, so we need to provide the values of $\langle \tau_h T, \phi \rangle$, $\langle d_\lambda T, \phi \rangle$ and $\langle \widetilde{T}, \phi \rangle$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$. If $T = T_f$ is the tempered distribution associated with $f \in L^1(\mathbb{R}^n)$, then for $g \in \mathscr{S}(\mathbb{R}^n)$, the change of variable formula implies that

$$\begin{aligned}
\langle \tau_h f, g \rangle &= \int_{\mathbb{R}^n} f(x-h)g(x)\,dx = \int_{\mathbb{R}^n} f(x)g(x+h)\,dx = \langle f, \tau_{-h} g \rangle, \\
\langle d_\lambda f, g \rangle &= \int_{\mathbb{R}^n} f(\lambda^{-1}x)g(x)\,dx = \int_{\mathbb{R}^n} f(x)g(\lambda x)\lambda^n\,dx = \langle f, \lambda^n d_{\lambda^{-1}} g \rangle, \\
\langle \widetilde{f}, g \rangle &= \int_{\mathbb{R}^n} f(-x)g(x)\,dx = \int_{\mathbb{R}^n} f(x)g(-x)\,dx = \langle f, \widetilde{g} \rangle.
\end{aligned}$$

The computations above motivate the following

Definition 9.46. Let $h \in \mathbb{R}^n$, $\lambda > 0$, and τ_h and d_λ be the translation and dilation operator given in Definition 9.45. For $T \in \mathscr{S}(\mathbb{R}^n)'$, $\tau_h T$, $d_\lambda T$ and $\widetilde{T}$ are the tempered distributions defined by

$$\langle \tau_h T, \phi \rangle = \langle T, \tau_{-h}\phi \rangle, \quad \langle d_\lambda T, \phi \rangle = \langle T, \lambda^n d_{\lambda^{-1}}\phi \rangle \text{ and } \langle \widetilde{T}, \phi \rangle = \langle T, \widetilde{\phi} \rangle \quad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).$$

We note that $\tau_h T$, $d_\lambda T$ and $\widetilde{T}$ are tempered distributions since

$$\begin{aligned}
p_k(\tau_{-h}\phi) &\leqslant \sup_{x\in\mathbb{R}^n, |\alpha|\leqslant k} \langle x \rangle^k \big| D^\alpha \phi(x-h) \big| \leqslant \sup_{x\in\mathbb{R}^n, |\alpha|\leqslant k} \langle x+h \rangle^k \big| D^\alpha \phi(x) \big| \\
&\leqslant \Big(\sup_{x\in\mathbb{R}^n} \frac{\langle x+h \rangle^2}{\langle x \rangle^2} \Big)^{\frac{k}{2}} p_k(\phi) \leqslant (1+|h|)^k p_k(\phi), \\
p_k(\lambda^n d_{\lambda^{-1}}\phi) &\leqslant \lambda^n \sup_{x\in\mathbb{R}^n, |\alpha|\leqslant k} \langle x \rangle^k \lambda^{|\alpha|} \big| (D^\alpha \phi)(\lambda x) \big| \leqslant \lambda^n \max\{\lambda^k, \lambda^{-k}\} p_k(\phi), \\
p_k(\widetilde{\phi}) &= p_k(\phi)
\end{aligned}$$

so that by the fact that $\big|\langle T, \phi \rangle\big| \leqslant C_k p_k(\phi)$ for $k \gg 1$, for all $\phi \in \mathscr{S}(\mathbb{R}^n)$ we have

$$\begin{aligned}
\big|\langle \tau_h T, \phi \rangle\big| &= \big|\langle T, \tau_{-h}\phi \rangle\big| \leqslant C_k (1+|h|)^k p_k(\phi) = \widetilde{C}_k p_k(\phi), \\
\big|\langle d_\lambda T, \phi \rangle\big| &= \big|\langle T, \lambda^n d_{\lambda^{-1}}\phi \rangle\big| \leqslant C_k \lambda^n \max\{\lambda^k, \lambda^{-k}\} p_k(\phi) = \widetilde{C}_k p_k(\phi), \\
\big|\langle \widetilde{T}, \phi \rangle\big| &= \big|\langle T, \widetilde{\phi} \rangle\big| \leqslant C_k p_k(\phi).
\end{aligned}$$

Example 9.47. Let $\omega, h \in \mathbb{R}^n$ and $\lambda > 0$.

1. $\tau_h \delta_\omega = \delta_{\omega+h}$ since if $\phi \in \mathscr{S}(\mathbb{R}^n)$, $\langle \tau_h \delta_\omega, \phi \rangle = \langle \delta_\omega, \tau_{-h}\phi \rangle = \phi(\omega+h) = \langle \delta_{\omega+h}, \phi \rangle$.

2. $d_\lambda \delta_\omega = \lambda^n \delta_{\lambda\omega}$ since if $\phi \in \mathscr{S}(\mathbb{R}^n)$, $\langle d_\lambda \delta_\omega, \phi\rangle = \langle \delta_\omega, \lambda^n d_{1/\lambda}\phi\rangle = \lambda^n \phi(\lambda\omega) = \langle \lambda^n \delta_{\lambda\omega}, \phi\rangle$.

3. $\widetilde{\delta_\omega} = \delta_{-\omega}$ since if $\phi \in \mathscr{S}(\mathbb{R}^n)$, $\langle \widetilde{\delta_\omega}, \phi\rangle = \langle \delta_\omega, \widetilde{\phi}\rangle = \phi(-\omega) = \langle \delta_{-\omega}, \phi\rangle$.

From the experience of defining the translation, dilation and reflection of tempered distributions, now we can talk about how to defined Fourier transform of tempered distributions. Recall that in Lemma 9.28 we have established that

$$\langle \widehat{f}, g\rangle = \langle f, \widehat{g}\rangle \quad \text{and} \quad \langle \check{f}, g\rangle = \langle f, \check{g}\rangle \qquad \forall\, f, g \in L^1(\mathbb{R}^n)\,.$$

Since the identities above hold for all L^1-functions f (and L^1-functions corresponds to tempered distributions T_f through (9.4.5)), we expect that the Fourier transform of tempered distributions has to satisfy the identities above as well. Let $T \in \mathscr{S}(\mathbb{R}^n)'$ be given, and define $\widehat{T} : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ by

$$\widehat{T}(\phi) = \langle \widehat{T}, \phi\rangle \equiv \langle T, \widehat{\phi}\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,. \tag{9.4.7}$$

Let $k \geqslant 2$ and k is a multiple of 4. Then

$$\begin{aligned}
p_k(\widehat{\phi}) &= \sup_{\xi\in\mathbb{R}^n, |\alpha|\leqslant k} \langle \xi\rangle^k \big|D^\alpha \widehat{\phi}(\xi)\big| = \sup_{\xi\in\mathbb{R}^n, |\alpha|\leqslant k} \langle \xi\rangle^k \Big|\mathscr{F}_x\big[x^\alpha \phi(x)\big](\xi)\Big| \\
&\leqslant \sup_{\xi\in\mathbb{R}^n, |\alpha|\leqslant k} (n+1)^{\frac{k}{2}-1}\big(1 + |\xi_1|^k + \cdots + |\xi_n|^k\big)\Big|\mathscr{F}_x\big[x^\alpha \phi(x)\big](\xi)\Big| \\
&\leqslant (n+1)^{\frac{k}{2}-1} \sup_{\xi\in\mathbb{R}^n, |\alpha|\leqslant k} \Big|\mathscr{F}_x\big[(1 + \partial_{x_1}^k + \cdots + \partial_{x_n}^k)\big(x^\alpha \phi(x)\big)\big](\xi)\Big|.
\end{aligned}$$

Since

$$\begin{aligned}
\sup_{\xi\in\mathbb{R}^n, |\alpha|\leqslant k} \Big|\mathscr{F}_x\big[x^\alpha \phi(x)\big](\xi)\Big| &\leqslant \sup_{|\alpha|\leqslant k} \int_{\mathbb{R}^n} \big|x^\alpha \phi(x)\big|\, dx \leqslant \int_{\mathbb{R}^n} \langle x\rangle^k \big|\phi(x)\big|\, dx \\
&\leqslant \big\|\langle \cdot\rangle^{-n-1}\big\|_{L^1(\mathbb{R}^n)} \sup_{x\in\mathbb{R}^n} \langle x\rangle^{n+k+1}\big|\phi(x)\big| \leqslant \big\|\langle \cdot\rangle^{-n-1}\big\|_{L^1(\mathbb{R}^n)}\, p_{n+k+1}(\phi)
\end{aligned}$$

and for $1 \leqslant j \leqslant n$,

$$\begin{aligned}
&\sup_{\xi\in\mathbb{R}^n, |\alpha|\leqslant k} \Big|\mathscr{F}_x\big[\partial_{x_j}^k (x^\alpha \phi(x)\big](\xi)\Big| \leqslant \sum_{\ell=0}^k C_\ell^k \sup_{\xi\in\mathbb{R}^n, |\alpha|\leqslant k} \Big|\mathscr{F}_x\big[(\partial_{x_j}^{k-\ell} x^\alpha)\partial_{x_j}^\ell \phi(x)\big](\xi)\Big| \\
&\qquad \leqslant \sum_{\ell=0}^k C_\ell^k \sup_{|\alpha|\leqslant k} \int_{\mathbb{R}^n} \big|(\partial_{x_j}^{k-\ell} x^\alpha)\partial_{x_j}^\ell \phi(x)\big|\, dx \leqslant \sum_{\ell=0}^k C_\ell^k \sup_{|\alpha|\leqslant k} |\alpha|!\big\|\langle \cdot\rangle^{|\alpha|-k+\ell}\partial_{x_j}^\ell \phi(\cdot)\big\|_{L^1(\mathbb{R}^n)} \\
&\qquad \leqslant \sum_{\ell=0}^k C_\ell^k k! \sup_{|\beta|=\ell} \big\|\langle \cdot\rangle^\ell D^\beta \phi(\cdot)\big\|_{L^1(\mathbb{R}^n)} \leqslant k! \sum_{\ell=0}^k C_\ell^k \big\|\langle \cdot\rangle^{-n-1}\big\|_{L^1(\mathbb{R}^n)} p_{n+\ell+1}(\phi) \\
&\qquad \leqslant k!\big\|\langle \cdot\rangle^{-n-1}\big\|_{L^1(\mathbb{R}^n)}\, p_{n+k+1}(\phi) \sum_{\ell=0}^k C_\ell^k = k! 2^k \big\|\langle \cdot\rangle^{-n-1}\big\|_{L^1(\mathbb{R}^n)}\, p_{n+k+1}(\phi)\,,
\end{aligned}$$

we conclude that

$$p_k(\widehat{\phi}) \leqslant (n+1)^{\frac{k}{2}-1}(1+nk!2^k)\big\|\langle\cdot\rangle^{-n-1}\big\|_{L^1(\mathbb{R}^n)}p_{n+k+1}(\phi) = \bar{C}(n,k)p_{n+k+1}(\phi). \tag{9.4.8}$$

Therefore,

$$\big|\langle\widehat{T},\phi\rangle\big| = \big|\langle T,\widehat{\phi}\rangle\big| \leqslant C_k p_k(\widehat{\phi}) \leqslant C_k\bar{C}(n,k)p_{k+n+1}(\phi) \qquad \forall\, k \gg 1 \tag{9.4.9}$$

which shows that $\widehat{T}$ defined by (9.4.7) is a tempered distribution. Similarly, $\widecheck{T} : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ defined by $\langle\widecheck{T},\phi\rangle = \langle T,\widecheck{\phi}\rangle$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$ is also a tempered distribution. The discussion above leads to the following

Definition 9.48. Let $T \in \mathscr{S}(\mathbb{R}^n)'$. The Fourier transform of T and the Fourier $*$ transform of T, denoted by $\widehat{T}$ and $\widecheck{T}$ respectively, are tempered distributions given by

$$\langle\widehat{T},\phi\rangle = \langle T,\widehat{\phi}\rangle \qquad \text{and} \qquad \langle\widecheck{T},\phi\rangle = \langle T,\widecheck{\phi}\rangle \qquad \forall\,\phi \in \mathscr{S}(\mathbb{R}^n).$$

In other words, if $T \in \mathscr{S}(\mathbb{R}^n)'$, then $\widehat{T},\widecheck{T} \in \mathscr{S}(\mathbb{R}^n)'$ as well and the actions of $\widehat{T}$, $\widecheck{T}$ on $\phi \in \mathscr{S}(\mathbb{R}^n)$ are given in the relations above.

Example 9.49 (The Fourier transform of the Dirac delta function)**.** Consider the Dirac delta function $\delta : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ defined in Example 9.43. Then for $\phi \in \mathscr{S}(\mathbb{R}^n)$,

$$\langle\delta,\widehat{\phi}\rangle = \widehat{\phi}(0) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\phi(x)e^{-ix\cdot 0}\,dx = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\phi(x)\,dx = \langle\frac{1}{\sqrt{2\pi}^n},\phi\rangle;$$

thus the Fourier transform of the Dirac delta function is a constant function and $\widehat{\delta}(\xi) = \dfrac{1}{\sqrt{2\pi}^n}$. Similarly, $\widecheck{\delta}(\xi) = \dfrac{1}{\sqrt{2\pi}^n}$, so $\widehat{\delta} = \widecheck{\delta}$.

Next we consider the Fourier transform of δ_ω, the Dirac delta function at point $\omega \in \mathbb{R}^n$. Note that for $\phi \in \mathscr{S}(\mathbb{R}^n)$,

$$\langle\delta_\omega,\widehat{\phi}\rangle = \widehat{\phi}(\omega) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\phi(x)e^{-ix\cdot\omega}\,dx = \Big\langle\frac{e^{-ix\cdot\omega}}{\sqrt{2\pi}^n},\phi\Big\rangle \equiv \langle\widehat{\delta_\omega},\phi\rangle;$$

thus the Fourier transform of the Dirac delta function at point ω is the function $\widehat{\delta_\omega}(\xi) = \dfrac{e^{-i\xi\cdot\omega}}{\sqrt{2\pi}^n}$. The inverse Fourier transform of δ_ω can be computed in the same fashion and we have $\widecheck{\delta_\omega}(\xi) = \dfrac{e^{i\xi\cdot\omega}}{\sqrt{2\pi}^n}$. We note that $\widecheck{\delta_\omega} = \widetilde{\widehat{\delta_\omega}} = \widehat{\widetilde{\delta_\omega}}$.

Symbolically, "assuming" that $\langle\delta_\omega,\phi\rangle = \phi(\omega)$ for all continuous function ϕ,

$$\widehat{\delta_\omega}(\xi) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\delta_\omega(x)e^{-ix\cdot\xi}\,dx = \frac{1}{\sqrt{2\pi}^n}e^{-ix\cdot\xi}\Big|_{x=\omega} = \frac{e^{-i\xi\cdot\omega}}{\sqrt{2\pi}^n}$$

and

$$\check{\delta_\omega}(\xi) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} \delta_\omega(x) e^{ix\cdot\xi}\,dx = \frac{1}{\sqrt{2\pi}^n} e^{ix\cdot\xi}\Big|_{x=\omega} = \frac{e^{i\xi\cdot\omega}}{\sqrt{2\pi}^n}\,.$$

Example 9.50 (The Fourier transform of $e^{ix\cdot\omega}$)**.** By "definition" and the Fourier inversion formula, for $\phi \in \mathscr{S}(\mathbb{R}^n)$ we have

$$\langle e^{ix\cdot\omega}, \widehat{\phi}\rangle = \int_{\mathbb{R}^n} e^{ix\cdot\omega}\widehat{\phi}(x)\,dx = \sqrt{2\pi}^n \cdot \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} \widehat{\phi}(x) e^{ix\cdot\omega}\,dx = \sqrt{2\pi}^n \check{\widehat{\phi}}(\omega) = \sqrt{2\pi}^n \phi(\omega)\,;$$

thus

$$\langle e^{ix\cdot\omega}, \widehat{\phi}\rangle = \sqrt{2\pi}^n \phi(\omega) = \langle \sqrt{2\pi}^n \delta_\omega, \phi\rangle.$$

Therefore, the Fourier transform of the function $s(x) = e^{ix\cdot\omega}$ is $\sqrt{2\pi}^n\delta_\omega$, where δ_ω is the Dirac delta function at point ω introduced in Example 9.49. We note that this result also implies that

$$\widehat{\check{\delta_\omega}} = \delta_\omega \qquad \forall\, \omega \in \mathbb{R}^n\,.$$

Similarly, $\check{\widehat{\delta_\omega}} = \delta_\omega$ for all $\omega \in \mathbb{R}^n$; thus the Fourier inversion formula is also valid for the Dirac δ function.

Example 9.51 (The Fourier Transform of the Sine function)**.** Let $s(x) = \sin\omega x$, where ω denotes the frequency of this sine wave. Since $\sin\omega x = \dfrac{e^{i\omega x} - e^{-i\omega x}}{2i}$, we conclude that the Fourier transform of $s(x) = \sin\omega x$ is

$$\frac{\sqrt{2\pi}}{2i}\big(\delta_\omega - \delta_{-\omega}\big)$$

since if T_1, T_2 are tempered distributions, then $T = T_1 + T_2$ satisfies

$$\langle \widehat{T}, \phi\rangle = \langle T_1 + T_2, \widehat{\phi}\rangle = \langle T_1, \widehat{\phi}\rangle + \langle T_2, \widehat{\phi}\rangle = \langle \widehat{T_1}, \phi\rangle + \langle \widehat{T_2}, \phi\rangle = \langle \widehat{T_1} + \widehat{T_2}, \phi\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)$$

which shows that $\widehat{T} = \widehat{T_1} + \widehat{T_2}$.

Theorem 9.52. *Let $T \in \mathscr{S}(\mathbb{R}^n)'$. Then $\check{\widehat{T}} = \widehat{\check{T}} = T$.*

Proof. To see that $\check{\widehat{T}}$ and T are the same tempered distribution, we need to show that $\langle \check{\widehat{T}}, \phi\rangle = \langle T, \phi\rangle$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$. Nevertheless, by the definition of the Fourier transform and the inverse Fourier transform of tempered distributions,

$$\langle \check{\widehat{T}}, \phi\rangle = \langle \widehat{T}, \check{\phi}\rangle = \langle T, \widehat{\check{\phi}}\rangle = \langle T, \phi\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

The identity $\widehat{\check{T}} = T$ can be proved in the same fashion. $\square$

Example 9.53 (The Fourier Transform of the sinc function)**.** The rect/rectangle function, also called the gate function or windows function, is a function $\Pi : \mathbb{R} \to \mathbb{R}$ defined by

$$\Pi(x) = \begin{cases} 1 & \text{if } |x| < 1\,, \\ 0 & \text{if } |x| \geqslant 1\,. \end{cases}$$

Since $\Pi \in L^1(\mathbb{R})$, we can compute its (inverse) Fourier transform in the usual way, and we have

$$\widehat{\Pi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \Pi(x) e^{-ix\xi}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-ix\xi}\, dx = \frac{1}{\sqrt{2\pi}} \frac{e^{-ix\xi}}{-i\xi}\Big|_{x=-1}^{x=1} = \sqrt{\frac{2}{\pi}} \frac{\sin \xi}{\xi} \qquad \forall\, \xi \neq 0$$

and $\widehat{\Pi}(0) = \sqrt{\dfrac{2}{\pi}}$. Define the ***unnormalized sinc function*** $\operatorname{sinc}(x) = \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0\,, \\ 1 & \text{if } x = 0\,. \end{cases}$

Then $\widehat{\Pi}(\xi) = \sqrt{\dfrac{2}{\pi}} \operatorname{sinc}(\xi)$. Similar computation shows that $\check{\Pi}(\xi) = \widehat{\Pi}(\xi) = \sqrt{\dfrac{2}{\pi}} \operatorname{sinc}(\xi)$.

Even though the sinc function is not integrable, we can apply Theorem 9.52 and see that

$$\widehat{\operatorname{sinc}}(\xi) = \widecheck{\operatorname{sinc}}(\xi) = \sqrt{\frac{\pi}{2}}\, \Pi(\xi) \qquad \forall\, \xi \in \mathbb{R}\,.$$

Theorem 9.54. *Let $T \in \mathscr{S}(\mathbb{R}^n)'$. Then*

$$\langle \widehat{\tau_h T}, \phi \rangle = \langle \widehat{T}(\xi), \phi(\xi) e^{-i\xi\cdot h} \rangle, \quad \langle \widehat{d_\lambda T}, \phi \rangle = \langle \widehat{T}, d_\lambda \phi \rangle \quad \textit{and} \quad \langle \widehat{\widehat{T}}, \phi \rangle = \langle \check{T}, \phi \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

A short-hand notation for identities above are $\widehat{\tau_h T}(\xi) = \widehat{T}(\xi) e^{-i\xi\cdot h}$, $\widehat{d_\lambda T}(\xi) = \lambda^n \widehat{T}(\lambda\xi)$, and $\widehat{\widehat{T}}(\xi) = \check{T}(\xi)$.

Proof. Let $\phi \in \mathscr{S}(\mathbb{R}^n)$. For $h \in \mathbb{R}^n$, define $\phi_h(x) = \phi(x) e^{-ix\cdot h}$. Then

$$(\tau_{-h}\widehat{\phi})(\xi) = \widehat{\phi}(\xi + h) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \phi(x) e^{-ix\cdot(\xi+h)}\, dx = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \phi(x) e^{-ix\cdot h} e^{-ix\cdot\xi}\, dx = \widehat{\phi_h}(\xi)\,.$$

By the definition of the Fourier transform of tempered distribution and the translation operator,

$$\langle \widehat{\tau_h T}, \phi \rangle = \langle T, \tau_{-h}\widehat{\phi} \rangle = \langle T, \widehat{\phi_h} \rangle = \langle \widehat{T}(x), \phi(x) e^{-ix\cdot h} \rangle = \langle \widehat{T}(\xi), \phi(\xi) e^{-i\xi\cdot h} \rangle.$$

On the other hand, for $\lambda > 0$,

$$(d_{\lambda^{-1}}\widehat{\phi})(\xi) = \widehat{\phi}(\lambda\xi) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \phi(x) e^{-ix\cdot(\lambda\xi)}\, dx = \lambda^{-n} \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \phi\big(\frac{x}{\lambda}\big) e^{-ix\cdot\xi}\, dx = \lambda^{-n} \widehat{d_\lambda \phi}(\xi)\,.$$

Therefore,

$$\langle \widehat{d_\lambda T}, \phi \rangle = \langle T, \lambda^n d_{\lambda^{-1}} \widehat{\phi} \rangle = \langle T, \widehat{d_\lambda \phi} \rangle = \langle \widehat{T}, d_\lambda \phi \rangle = \langle \lambda^n d_{\lambda^{-1}} \widehat{T}, \phi \rangle.$$

The identity $\langle \widehat{\widetilde{T}}, \phi \rangle = \langle \check{T}, \phi \rangle$ follows from that $\widetilde{\widehat{\phi}} = \check{\phi}$, and the detail proof is left to the readers. □

Remark 9.55. One can check easily that $\widehat{\tau_h f}(\xi) = \widehat{f}(\xi) e^{-i\xi\cdot h}$ and $\widehat{d_\lambda f}(\xi) = \lambda^n \widehat{f}(\lambda\xi)$ if $f \in L^1(\mathbb{R}^n)$.

Next we define the convolution of a tempered distribution and a Schwartz function. Before proceeding, we note that if $f, g, \phi \in \mathscr{S}(\mathbb{R}^n)$, then the Fubini Theorem implies that

$$\begin{aligned}
\langle f * g, \phi \rangle &= \int_{\mathbb{R}^n} (f * g)(x)\phi(x)\, dx = \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} f(y) g(x-y)\, dy \Big) \phi(x)\, dx \\
&= \int_{\mathbb{R}^n} f(y) \Big(\int_{\mathbb{R}^n} g(x-y)\phi(x)\, dx \Big)\, dy \\
&= \int_{\mathbb{R}^n} f(y) \Big(\int_{\mathbb{R}^n} \widetilde{g}(y-x)\phi(x)\, dx \Big)\, dy = \langle f, \widetilde{g} * \phi \rangle.
\end{aligned}$$

The change of variable formula further shows that

$$\begin{aligned}
(\widetilde{g} * \phi)(y) &= \Big(\int_{\mathbb{R}^n} \widetilde{g}(x)\phi(y-x)\, dx = \int_{\mathbb{R}^n} \widetilde{g}(-x)\phi(y+x)\, dx \\
&= \int_{\mathbb{R}^n} g(x)\widetilde{\phi}(-y-x)\, dx = (g * \widetilde{\phi})(-y) = \widetilde{g * \widetilde{\phi}}(y);
\end{aligned}$$

thus

$$\langle f * g, \phi \rangle = \langle f, \widetilde{g} * \phi \rangle = \langle f, \widetilde{g * \widetilde{\phi}} \rangle = \langle \widetilde{f}, g * \widetilde{\phi} \rangle \qquad \forall\, f, g, \phi \in \mathscr{S}(\mathbb{R}^n).$$

The identity above serves as the origin of the convolution of a tempered distribution and a Schwartz function.

Definition 9.56 (Convolution)**.** Let $T \in \mathscr{S}(\mathbb{R}^n)'$ and $g \in \mathscr{S}(\mathbb{R}^n)$. The convolution of T and g, denoted by $T * g$, is the tempered distribution given by

$$\langle T * g, \phi \rangle = \langle T, \widetilde{g} * \phi \rangle = \langle \widetilde{T}, g * \widetilde{\phi} \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n),$$

where $\widetilde{T}$ is the tempered distribution given in Definition 9.46.

Remark 9.57. We will explain why $T * g \in \mathscr{S}(\mathbb{R}^n)'$ (if $T \in \mathscr{S}(\mathbb{R}^n)'$ and $g \in \mathscr{S}(\mathbb{R}^n)$) later in Remark 9.60. For the time being we can temporarily treat the convolution given above as a "computational" definition (without knowing that if $T * g$ is continuous); that is, $T * g : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ is defined by

$$(T * g)(\phi) \equiv \langle T, \widetilde{g} * \phi \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)$$

since $\langle T, \widetilde{g} * \phi \rangle$ is well-defined, and $\langle T * g, \phi \rangle$ is another expression of $(T * g)(\phi)$.

Example 9.58. Let δ_ω be the Dirac delta function at point $\omega \in \mathbb{R}^n$, and $g \in \mathscr{S}(\mathbb{R}^n)$. Then $\delta_\omega * g = \tau_\omega g$ since if $\phi \in \mathscr{S}(\mathbb{R}^n)$,

$$\langle \delta_\omega, \widetilde{g} * \phi \rangle = (\widetilde{g} * \phi)(\omega) = \int_{\mathbb{R}^n} \widetilde{g}(y)\phi(\omega - y)\, dy = \int_{\mathbb{R}^n} g(z - \omega)\phi(z)\, dz = \big\langle \tau_\omega g, \phi \big\rangle$$

In symbol,

$$(\delta_\omega * g)(x) = \int_{\mathbb{R}^n} \delta_\omega(y) g(x - y)\, dy = g(x - \omega) = (\tau_\omega g)(x)\,. \tag{9.4.10}$$

Similar to Theorem 9.26 and Corollary 9.27, the product and the convolutions of functions are related under Fourier transform.

Theorem 9.59. *Let $T \in \mathscr{S}(\mathbb{R}^n)'$ and $g \in \mathscr{S}(\mathbb{R}^n)$. Then*

$$\langle T \star g, \widehat{\phi} \rangle = \langle \widehat{T}, \widehat{g}\phi \rangle \quad \textit{and} \quad \langle T \star g, \check{\phi} \rangle = \langle \check{T}, \check{g}\phi \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,,$$

and

$$\langle \widehat{T} \star \widehat{g}, \phi \rangle = \langle T, g\widehat{\phi} \rangle \quad \textit{and} \quad \langle \check{T} \star \check{g}, \phi \rangle = \langle T, g\check{\phi} \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,,$$

where $S \star h = \dfrac{1}{\sqrt{2\pi}^{\,n}}(S * h)$ if $S \in \mathscr{S}(\mathbb{R}^n)'$ and $h \in \mathscr{S}(\mathbb{R}^n)$.

Proof. Note that the "convolution" $\star$ also satisfies that

$$\langle T \star g, \phi \rangle = \langle T, \widetilde{g} \star \phi \rangle = \langle \widetilde{T}, g \star \widetilde{\phi} \rangle.$$

By Theorem 9.26 and Corollary 9.27,

$$\langle T \star g, \widehat{\phi} \rangle = \langle T, \widetilde{g} \star \widehat{\phi} \rangle = \langle \check{\widehat{T}}, \widetilde{g} \star \widehat{\phi} \rangle = \langle \widehat{T}, \mathscr{F}^*[\widetilde{g} \star \widehat{\phi}] \rangle = \big\langle \widehat{T}, \check{\widetilde{g}}\check{\widehat{\phi}} \big\rangle = \langle \widehat{T}, \widehat{g}\phi \rangle$$

and by the definition of the convolution of tempered distributions and Schwartz functions,

$$\langle T, g\widehat{\phi} \rangle = \langle \check{\widehat{T}}, g\widehat{\phi} \rangle = \langle \widehat{T}, \mathscr{F}^*(g\widehat{\phi}) \rangle = \langle \widehat{T}, \check{g} \star \phi \rangle = \langle \widehat{T}, \widetilde{\widehat{g}} \star \phi \rangle = \langle \widehat{T} \star \widehat{g}, \phi \rangle.$$

The counterpart for the inverse Fourier transform can be proved similarly. $\square$

Remark 9.60. Let $g, \phi \in \mathscr{S}(\mathbb{R}^n)$, and $T \in \mathscr{S}(\mathbb{R}^n)'$ satisfy $\big|\langle T, u\rangle\big| \leqslant C_k p_k(u)$ for all $u \in \mathscr{S}(\mathbb{R}^n)$ and $k \gg 1$. By Theorem 9.59, we find that

$$\langle T * g, \phi\rangle = \sqrt{2\pi}^{\,n}\langle T \star g, \hat{\check{\phi}}\rangle = \sqrt{2\pi}^{\,n}\langle \widehat{T}, \widehat{g}\check{\phi}\rangle.$$

By the fact that

$$\begin{aligned} p_k(gh) &= \sup_{x\in\mathbb{R}^n, |\alpha|\leqslant k} \langle x\rangle^k \big|D^\alpha(gh)(x)\big| \leqslant \sup_{x\in\mathbb{R}^n, |\alpha|\leqslant k} \sum_{0\leqslant\beta\leqslant\alpha} C^\alpha_\beta \langle x\rangle^k \big|D^{\alpha-\beta}g(x)D^\beta h(x)\big| \\ &\leqslant \Big(\sup_{|\alpha|\leqslant k}\sum_{0\leqslant\beta\leqslant\alpha} C^\alpha_\beta\Big) p_k(g)p_k(h) \equiv M_k p_k(g)p_k(h) \qquad \forall\, g, h \in \mathscr{S}(\mathbb{R}^n), \end{aligned}$$

we conclude from (9.4.8) and (9.4.9) that for $k \gg 1$,

$$\begin{aligned} \big|\langle T * g, \phi\rangle\big| &\leqslant \sqrt{2\pi}^{\,n} C_k \bar{C}(n,k) p_{k+n+1}\big(\widehat{g}\check{\phi}\big) \leqslant \sqrt{2\pi}^{\,n} C_k \bar{C}(n,k) M_k p_{k+n+1}\big(\widehat{g}\big) p_{k+n+1}\big(\hat{\check{\phi}}\big) \\ &\leqslant \sqrt{2\pi}^{\,n} C_k M_k \bar{C}(n,k)\bar{C}(n,k+n+1)^2 p_{k+2n+2}(g) p_{k+2n+2}\big(\widetilde{\phi}\big) \\ &= \widetilde{C}(n,k) p_{k+2n+2}(g) p_{k+2n+2}(\phi). \end{aligned}$$

Therefore, $T * g$ is a tempered distribution.

Remark 9.61. For $T \in \mathscr{S}(\mathbb{R}^n)'$ and $g \in \mathscr{S}(\mathbb{R}^n)$, define $gT : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ by

$$\langle gT, \phi\rangle = \langle T, g\phi\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).$$

Then the fact that $T * g \in \mathscr{S}(\mathbb{R}^n)'$ (from Remark 9.60) and Theorem 9.59 show that

$$\langle \widehat{T \star g}, \phi\rangle = \langle \widehat{g}\widehat{T}, \phi\rangle \quad \text{and} \quad \langle \widecheck{T \star g}, \phi\rangle = \langle \check{g}\check{T}, \phi\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n),$$

and

$$\langle \widehat{T} \star \widehat{g}, \phi\rangle = \langle \widehat{gT}, \phi\rangle \quad \text{and} \quad \langle \check{T} \star \check{g}, \phi\rangle = \langle \widecheck{gT}, \phi\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n),$$

In other words, we have $\widehat{T \star g} = \widehat{g}\widehat{T}$, $\widecheck{T \star g} = \check{g}\check{T}$, $\widehat{gT} = \widehat{T} \star \widehat{g}$ and $\widecheck{gT} = \check{T} \star \check{g}$ in $\mathscr{S}(\mathbb{R}^n)'$. Therefore, Theorem 9.59 can be viewed as the generalization of Theorem 9.26 and Corollary 9.27.

Remark 9.62. If $S \in \mathscr{S}(\mathbb{R}^n)'$ satisfies that $S * \phi \in \mathscr{S}(\mathbb{R}^n)$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$, we can also define the convolution of T and S by

$$\langle T * S, \phi\rangle = \langle \widetilde{T}, S * \widetilde{\phi}\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).$$

In other words, it is possible to define the convolution of two tempered distributions.

For example, from Example 9.58 we find that $\delta_\omega * \phi = \tau_\omega \phi$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$; thus $\delta_\omega * \phi \in \mathscr{S}(\mathbb{R}^n)$ for all $\mathscr{S}(\mathbb{R}^n)$ (and $\omega \in \mathbb{R}^n$). Therefore, if T is a tempered distribution, $T * \delta_\omega$ is also a tempered distribution and is given by

$$\langle T * \delta_\omega, \phi \rangle = \big\langle \widetilde{T}, \tau_\omega \widetilde{\phi} \big\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).$$

Further computation shows that

$$\langle T * \delta_\omega, \phi \rangle = \big\langle \widetilde{T}, \widetilde{\tau_{-\omega}\phi} \big\rangle = \big\langle T, \tau_{-\omega}\phi \big\rangle = \big\langle \tau_\omega T, \phi \big\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).$$

The identity above shows that $T * \delta_\omega = \tau_\omega T$ for all $T \in \mathscr{S}(\mathbb{R}^n)'$. This formula agrees with (9.4.10).

9.5 Exercises

§ 9.1 The Definition of the Fourier Transform

Problem 9.1. Find the Fourier transform of the following functions.

1. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = xe^{-tx^2}$ for $t > 0$.
2. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \mathbf{1}_{(-a,a)}(x)$, the characteristic (indicator) function of the set $(-a, a)$.
3. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \begin{cases} e^{-tx} & \text{if } x > 0\,, \\ 0 & \text{if } x \leqslant 0\,, \end{cases}$ where $t > 0$.

Problem 9.2. Compute the Fourier transform of the following functions.

(a) $y = x\mathbf{1}_{[-1,1]}(x)$;

(b) $y = \sin x \cdot \mathbf{1}_{[-\pi,\pi]}(x)$;

(c) $y = e^{-x}H(x)$;

(d) $y = e^{-|x|}\cos x$;

(e) $y = \dfrac{1}{x^2 + 6x + 13}$;

(f) $y = \dfrac{x}{(x^2+1)^2}$;

(g) $y = \dfrac{1}{(x^2+1)^2}$.

Here, H is the Heaviside function (that is, $H(x) = \mathbf{1}_{[1,\infty)}(x)$) and $\mathbf{1}_{[a,b]}(x)$ is the characteristic function of $[a,b]$; that is, $\mathbf{1}_{[a,b]}(x) = 1$ for $x \in [a,b]$ and 0 else.

Problem 9.3. Suppose that the Fourier transform of a Schwartz function $f \in \mathscr{S}(\mathbb{R})$ is $\widehat{f}(\xi)$. Find the Fourier transform of the function $y = f(2x+1)\cos x$.

Problem 9.4. Let $\alpha > 0$ be given. Show that the Fourier transform of the function

$$f(x) = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-t} e^{-\frac{|x|^2}{t}}\, dt\,, \qquad x \in \mathbb{R}^n\,,$$

is positive.

§ 9.2 Some Further Properties of the Fourier Transform

Problem 9.5. Suppose that the Fourier transform of a Schwartz function $f \in \mathscr{S}(\mathbb{R})$ is $\widehat{f}(\xi)$. Find the Fourier transform of the function $y = f(2x+1)\cos x$.

§ 9.3 The Fourier Inversion Formula

Problem 9.6. Prove Corollary 9.27.

Problem 9.7. Suppose that $f : \mathbb{R} \to \mathbb{C}$ is continuous, absolutely integrable, and $\widehat{f}(\xi) = \dfrac{\ln(1+\xi^2)}{\xi^2}$. Find $f(0)$ and $\displaystyle\int_{-\infty}^{\infty} f(x)\, dx$.

Problem 9.8. 1. Let $f : \mathbb{R} \to \mathbb{C}$ be a continuous integrable function such that $\widehat{f}$ is also integrable. Show that

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(y) \cos[(x-y)\xi]\, dy \Big)\, d\xi \qquad \forall\, x \in \mathbb{R}\,.$$

2. If in addition to condition in 1, f is an even function. Show that

$$f(x) = \frac{2}{\pi} \int_0^\infty \Big(\int_0^\infty f(y) \cos(x\xi)\cos(y\xi)\, dy \Big)\, d\xi\,.$$

3. If in addition to condition in 1, f is an odd function. Show that

$$f(x) = \frac{2}{\pi} \int_0^\infty \Big(\int_0^\infty f(y) \sin(x\xi)\sin(y\xi)\, dy \Big)\, d\xi\,.$$

4. For a function $g : [0, \infty) \to \mathbb{C}$ satisfying $\displaystyle\int_0^\infty |g(x)|\, dx < \infty$, the Fourier cosine transform and the Fourier sine transform of g, denoted by $\mathscr{F}_{\cos}[g]$ and $\mathscr{F}_{\sin}[g]$ respectively, are functions defined by

$$\mathscr{F}_{\cos}[g](\xi) = \sqrt{\frac{2}{\pi}} \int_0^\infty g(y) \cos(y\xi)\, dy \quad \text{and} \quad \mathscr{F}_{\sin}[g](\xi) = \sqrt{\frac{2}{\pi}} \int_0^\infty g(y) \sin(y\xi)\, dy\,.$$

(a) Show that if $\mathscr{F}_{\cos}[g] \in L^1(\mathbb{R})$, then

$$g(x) = \mathscr{F}_{\cos}\big[\mathscr{F}_{\cos}[g]\big](x) \qquad \text{whenever } x \in [0, \infty) \text{ and } g \text{ is continuous at } x,$$

or equivalently,

$$g(x) = \frac{2}{\pi} \int_0^\infty \Big(\int_0^\infty g(y) \cos(y\xi)\, dy \Big) \cos(x\xi)\, d\xi$$

whenever $x \in [0, \infty)$ and g is continuous at x.

(b) Show that if $\mathscr{F}_{\sin}[g] \in L^1(\mathbb{R})$, then

$$g(x) = \mathscr{F}_{\sin}\big[\mathscr{F}_{\sin}[g]\big](x) \qquad \text{whenever } x \in [0, \infty) \text{ and } g \text{ is continuous at } x,$$

or equivalently,

$$g(x) = \frac{2}{\pi} \int_0^\infty \Big(\int_0^\infty g(y) \sin(y\xi)\, dy \Big) \sin(x\xi)\, d\xi$$

whenever $x \in (0, \infty)$ and g is continuous at x.

Hint of 4: Consider the even or odd extension of g, and apply conclusions in 2 and 3.

Problem 9.9. Let $f : \mathbb{R} \to \mathbb{C}$ be an integrable function, and $g(x) = f(x)\cos(ax)$, $h(x) = f(x)\sin(ax)$, where $a \neq 0$ is a real constant.

1. Show that $\widehat{g}(\xi) = \dfrac{1}{2}\big[\widehat{f}(\xi - a) + \widehat{f}(\xi + a)\big]$.

2. Show that $\mathscr{F}_{\cos}[g](\xi) = \dfrac{1}{2}\big[\mathscr{F}_{\cos}[f](\xi + a) + \mathscr{F}_{\cos}[f](\xi - a)\big]$.

3. Show that $\mathscr{F}_{\sin}[g](\xi) = \dfrac{1}{2}\big[\mathscr{F}_{\sin}[f](\xi + a) + \mathscr{F}_{\sin}[f](\xi - a)\big]$.

4. Show that $\mathscr{F}_{\cos}[h](\xi) = \dfrac{1}{2}\big[\mathscr{F}_{\cos}[f](\xi + a) - \mathscr{F}_{\cos}[f](\xi - a)\big]$.

5. Show that $\mathscr{F}_{\sin}[h](\xi) = \dfrac{1}{2}\big[\mathscr{F}_{\sin}[f](\xi - a) - \mathscr{F}_{\sin}[f](\xi + a)\big]$.

where $\mathscr{F}_{\cos}$ and $\mathscr{F}_{\sin}$ denote the Fourier cosine transform and sine transform, respectively.

Problem 9.10. Find the Fourier Sine and Cosine transform of the function $f(x) = e^{-x}$ and hence show that

$$\int_0^\infty \frac{\cos(x\xi)}{x^2+1}\,dx = \int_0^\infty \frac{x\sin(x\xi)}{x^2+1}\,dx = \frac{\pi}{2}e^{-\xi}\,.$$

Problem 9.11. Let $a > 0$. Find the Fourier sine transform of the function $f(x) = \dfrac{e^{-ax}}{x}$ and use it to evaluate

$$\int_0^\infty \sin x \arctan\frac{x}{a}\,dx\,.$$

Problem 9.12. Find the Fourier cosine integral representation of the function

$$f(x) = \begin{cases} x^2 & \text{if } 0 \leqslant x \leqslant a\,, \\ 0 & \text{if } x > a\,. \end{cases}$$

Problem 9.13. Let $f : [0,\infty) \to \mathbb{R}$ be defined by $f(x) = \begin{cases} \sin x & \text{if } 0 \leqslant x \leqslant \pi\,, \\ 0 & \text{if } x > \pi\,. \end{cases}$ Show that

$$f(x) = \frac{1}{\pi}\int_0^\infty \int_0^\infty \frac{2\cos(\xi x) + \cos\xi(\pi + x) + \cos\xi(\pi - x)}{1 - \xi^2}\,d\xi$$

and hence evaluate $\displaystyle\int_0^\infty \frac{\cos\frac{\pi t}{2}}{1-t^2}\,dt$.

Hint: Use the Fourier cosine integral representation of f.

Problem 9.14. Find the function f is its Fourier cosine transform is given by

(a) $\dfrac{\sin(a\xi)}{\xi}$. (b) $\dfrac{\xi}{1+\xi^2}$.

Problem 9.15. Find the function f is its Fourier sine transform is given by

(a) $e^{-a\xi}$, $a > 0$. (b) $\begin{cases} \dfrac{1}{\sqrt{2\pi}}\Big(a - \dfrac{\xi}{2}\Big) & \text{if } 0 \leqslant \xi \leqslant 2a\,, \\ 0 & \text{if } \xi > 2a\,. \end{cases}$

Problem 9.16. Solve the integral equation $\displaystyle\int_0^\infty f(x)\cos(\lambda x)\,dx = \begin{cases} 1 - \lambda & \text{if } 0 \leqslant \lambda \leqslant 1\,, \\ 0 & \text{if } \lambda > 1\,. \end{cases}$

Hence deduce that $\displaystyle\int_0^\infty \frac{\sin^2 t}{t^2}\,dt = \frac{\pi}{2}$.

Problem 9.17. A vector-valued function $\boldsymbol{u} = (u_1, u_2, \cdots, u_n) : \mathbb{R}^n \to \mathbb{R}^n$ is called a Schwartz function, still denoted by $\boldsymbol{u} \in \mathscr{S}(\mathbb{R}^n)$, if $u_j \in \mathscr{S}(\mathbb{R}^n)$ for all $1 \leqslant j \leqslant n$. Show the Korn inequality

$$\sum_{i,j=1}^{n} \big\|\epsilon_{ij}(\boldsymbol{u})\big\|_{L^2(\mathbb{R}^n)}^2 \geqslant \frac{1}{2} \sum_{i,j=1}^{n} \Big\|\frac{\partial u_j}{\partial x_i}\Big\|_{L^2(\mathbb{R}^n)}^2 \qquad \forall\, \boldsymbol{u} \in \mathscr{S}(\mathbb{R}^n)\,,$$

where $\epsilon_{ij}(\boldsymbol{u}) = \dfrac{1}{2}\Big(\dfrac{\partial u_i}{\partial x_j} + \dfrac{\partial u_j}{\partial x_i}\Big)$ is the symmetric part of $D\boldsymbol{u}$.

Hint: Use the Plancherel formula.

Problem 9.18. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a smooth function with compact support; that is, $\big\{x \in \mathbb{R}^n \,\big|\, \phi(x) \neq 0\big\}$ is bounded. Show that if $f \in L^1_{\rm loc}(\mathbb{R}^n)$, then the convolution $\phi * f$ is smooth and

$$D^\alpha(\phi * f)(x) = \big[(D^\alpha\phi) * f\big](x) = \int_{\mathbb{R}^n} (D^\alpha\phi)(x-y) f(y)\, dy\,,$$

where $\alpha = (\alpha_1, \cdots, \alpha_n)$ be a multi-index.

Note that the standard mollifiers $\{\eta_\varepsilon\}_{\varepsilon>0}$ are one of such kind of functions, so this problem shows that $\eta_\varepsilon * f$ is smooth if $f \in L^1_{\rm loc}(\mathbb{R}^n)$.

Problem 9.19. Find the Fourier transform of the function $f(x) = \begin{cases} 1 - x^2 & \text{if } |x| \leqslant 1\,, \\ 0 & \text{if } |x| > 1\,, \end{cases}$ and hence evaluate

$$\int_0^\infty \frac{x\cos x - \sin x}{x^3} \cos\frac{x}{2}\, dx\,.$$

Problem 9.20. Find the Fourier transform of the function $f(x) = \begin{cases} a - |x| & \text{if } |x| \leqslant a\,, \\ 0 & \text{if } |x| > a\,, \end{cases}$ and hence prove that

$$\int_0^\infty \frac{\sin^2 x}{x^2}\, dx = \frac{\pi}{2}\,.$$

Problem 9.21.
1. Let d_r denote the dilation operator defined by $d_r f(x) = f\big(\frac{x}{r}\big)$. Show that

$$\mathscr{F}(d_r f) = r^n d_{1/r}\mathscr{F}(f) \qquad \forall\, f \in \mathscr{S}(\mathbb{R}^n)\,. \tag{9.5.1}$$

2. In some occasions (especially in engineering applications), the Fourier transform and inverse Fourier transform of a (Schwartz) function f are defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-i2\pi x\cdot\xi} dx \qquad \text{and} \qquad \check{f}(x) = \int_{\mathbb{R}^n} f(\xi) e^{i2\pi x\cdot\xi} d\xi\,.$$

Show that under this definition, $\check{\hat{f}} = \hat{\check{f}} = f$ for all $f \in \mathscr{S}(\mathbb{R}^n)$. Note that you can use the Fourier Inversion Formula that we derive in class.

§ 9.4 The Fourier Transform of Generalized Functions

In the following problems, assume that you can use the Fourier inversion formula (without checking that whether f and $\hat{f}$ belong to $L^1(\mathbb{R}^n)$).

Problem 9.22. Let $f : \mathbb{R} \to \mathbb{R}$ be given by

$$f(x) = \int_0^2 \frac{\sqrt{t}}{1+t} e^{ixt} \, dt \,.$$

Compute the following.

(a) The Fourier transform of f. (b) $\displaystyle\int_{-\infty}^{\infty} f(x) \cos x \, dx$. (c) $\displaystyle\int_{-\infty}^{\infty} |f(x)|^2 \, dx$.

Problem 9.23. Suppose that $f : \mathbb{R} \to \mathbb{C}$ has the Fourier transform $\hat{f}(\xi) = \dfrac{1 - i\xi}{1 + i\xi} \dfrac{\sin \xi}{\xi}$. Compute $\displaystyle\int_{-\infty}^{\infty} |f(x)|^2 \, dx$.

Problem 9.24. Let function $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = \displaystyle\int_0^1 \sqrt{t} e^{t^2} \cos(xt) \, dt$. Compute $\displaystyle\int_{-\infty}^{\infty} \left|f'(x)\right|^2 dx$.

Problem 9.25. Use the Fourier transform to compute $\displaystyle\int_{-\infty}^{\infty} \frac{\sin x}{x(x^2+1)} \, dx$.

Problem 9.26. Let $f(x) = e^{-s|x|^2}$ and $g(x) = e^{-t|x|^2}$. Find the Fourier transform of f (and g) and use the inversion formula to compute $f * g$.

Problem 9.27. Use the Fourier transform to find a function f such that

$$\int_{-\infty}^{\infty} f(x-y) e^{-|y|} dy = 2e^{-|x|} - e^{-2|x|} \,.$$

Problem 9.28. For a and b positive, use the Fourier transform to compute the integrals

(a) $\displaystyle\int_{-\infty}^{\infty} \frac{\sin(at)\sin(b(u-t))}{t(u-t)} \, dt$. (b) $\displaystyle\int_{-\infty}^{\infty} \frac{\sin(at)\sin(bt)}{t^2} \, dt$.

Problem 9.29. Suppose that $f : \mathbb{R} \to \mathbb{C}$ has the Fourier transform $\hat{f}(\xi) = \dfrac{1}{|\xi|^3 + 1}$. Compute $\displaystyle\int_{-\infty}^{\infty} |(f * f')(x)|^2 \, dx$.

Problem 9.30. Let $a > 0$. Use the Fourier transforms of sinc and sinc^2, together with the basic tools of Fourier transform theory, such as Plancherel's identity, substitution, and etc. to show the following. (Use only rules from Fourier transform theory. You should not do any detailed computation such as integration by parts.)

(a) $\displaystyle\int_{-\infty}^{\infty} \frac{\sin^3(ax)}{x^3}\,dx = \frac{3a^2\pi}{4}$. (b) $\displaystyle\int_{-\infty}^{\infty} \frac{\sin^4(ax)}{x^4}\,dx = \frac{2a^3\pi}{3}$.

Problem 9.31. Deduce what you can about the Fourier transform $\widehat{f}(\xi)$ if you know that f satisfies the dilation equation

$$f(t) = f(2t) + f(2t-1) \qquad \forall\, t \in \mathbb{R}\,.$$

Problem 9.32. Show that every polynomial is a tempered distribution.

Problem 9.33. Let $\{\eta_\epsilon\}_{\epsilon>0}$ is the standard mollifiers, and $\phi \in \mathscr{S}(\mathbb{R}^n)$. Show that $\{\eta_\epsilon * \phi\}_{\epsilon>0}$ converges to ϕ in $\mathscr{S}(\mathbb{R}^n)$.

Problem 9.34. In this problem we consider the concept of the convergence of sequence of tempered distribution given by the following

Definition 9.63 (**Convergence in $\mathscr{S}(\mathbb{R}^n)'$**)**.** A sequence of distributions $T_n \in \mathscr{S}(\mathbb{R}^n)'$ is said to converge to $T \in \mathscr{S}(\mathbb{R}^n)'$ *in the sense of distribution*, or *in the distributional sense*, if $\langle T_n, \varphi\rangle \to \langle T, \varphi\rangle$ as $n \to \infty$ for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$.

Complete the following.

1. Show that if $T \in \mathscr{S}(\mathbb{R}^n)'$ and $\{\phi_n\}_{n=1}^{\infty} \subseteq \mathscr{S}(\mathbb{R}^n)$ is a sequence which converges to ϕ in $\mathscr{S}(\mathbb{R}^n)$, then $\lim\limits_{n\to\infty}\langle T, \phi_n\rangle = \langle T, \phi\rangle$.
2. Given the definition above, show that if $T \in \mathscr{S}(\mathbb{R}^n)'$, then $\{\eta_\epsilon * T\}_{\epsilon>0}$ converges to T in the sense of distribution, where $\{\eta_\epsilon\}_{\epsilon>0}$ is the standard mollifiers.

Problem 9.35. 1. Let $\alpha > 0$. Compute the Fourier transform of the function

$$f_\alpha(x) = \begin{cases} e^{-\alpha x} & \text{if } x \geqslant 0\,, \\ -e^{\alpha x} & \text{if } x < 0\,. \end{cases}$$

2. Show that $\lim\limits_{\alpha\to 0^+} \widehat{f}_\alpha(\xi) = \widehat{\text{sgn}}(\xi)$; that is,

$$\lim_{\alpha\to 0^+} \langle \widehat{f}_\alpha, \phi\rangle = \langle \widehat{\text{sgn}}(\xi), \phi\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,,$$

where sgn is the sign function given by

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x > 0\,, \\ -1 & \text{if } x < 0\,, \\ 0 & \text{if } x = 0\,. \end{cases} \tag{9.5.2}$$

3. Find the Fourier transform of the Heaviside function H using $H = \frac{1}{2}(1 + \operatorname{sgn})$.

4. Use the Fourier transform of sgn to compute the Fourier transform of the following functions:

   (a) $f_1(t) = \dfrac{1}{t}$. (b) $f_2(t) = |t|$.

   Use (a) and (b) to compute the Fourier transform of

   (c) $f_3(t) = -\dfrac{1}{t^2}$. (d) $f_4(t) = \dfrac{2}{t^3}$.

Problem 9.36. In this problem we discuss the derivative of tempered distributions. Complete the following.

1. Show that
$$\Big\langle \frac{\partial f}{\partial x_j}, g \Big\rangle = -\Big\langle f, \frac{\partial g}{\partial x_j} \Big\rangle \qquad \forall\, f, g \in \mathscr{S}(\mathbb{R}^n)\,.$$
   This leads to the definition of the derivatives of tempered distributions: Let $T \in \mathscr{S}(\mathbb{R}^n)'$ be a tempered distribution. The partial derivative of T w.r.t. x_j, denoted by $\dfrac{\partial T}{\partial x_j}$, is a tempered distribution defined by
$$\Big\langle \frac{\partial T}{\partial x_j}, \phi \Big\rangle = -\Big\langle T, \frac{\partial \phi}{\partial x_j} \Big\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,.$$
   Verify that $\dfrac{\partial T}{\partial x_j}$ is indeed a tempered distribution; that is, show that there exists a sequence $\{C_k\}_{k=1}^{\infty}$ such that
$$\Big|\Big\langle \frac{\partial T}{\partial x_j}, \phi \Big\rangle\Big| \leqslant C_k p_k(\phi) \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n) \text{ and } k \gg 1\,.$$

2. Show that for $1 \leqslant j \leqslant n$,
$$\mathscr{F}_x\Big[\frac{\partial T}{\partial x_j}\Big](\xi) = i\xi_j \widehat{T}(\xi) \qquad \text{and} \qquad \frac{\partial}{\partial x_j}\widehat{T}(\xi) = -i\mathscr{F}_x[xT(x)](\xi)$$

or to be more precise,

$$\Big\langle \widehat{\frac{\partial T}{\partial x_j}}, \phi \Big\rangle = \big\langle \widehat{T}(\xi), i\xi_j \phi(\xi) \big\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)$$

and

$$\Big\langle \frac{\partial}{\partial \xi_j} \widehat{T}(\xi), \phi(\xi) \Big\rangle = \big\langle T(x), -ix\widehat{\phi}(x) \big\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n)\,.$$

In other words, the Fourier transform of derivatives of tempered distributions still obeys Lemma 9.9 and 9.11.

Problem 9.37. Let $\{\eta_\epsilon\}_{\epsilon>0}$ be the standard mollifiers, and $T \in \mathscr{S}(\mathbb{R}^n)'$.

1. Show that the distribution $T * \eta_\epsilon$ is indeed a smooth function.

2. Show that if $T' = 0$ in $\mathscr{S}(\mathbb{R})'$, where the derivative of tempered distribution is given by Problem 9.36, then T is a constant; that is, there exists $C \in \mathbb{R}$ such that

$$\langle T, \phi \rangle = \langle C, \phi \rangle = C \int_{\mathbb{R}} \phi(x)\, dx \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,.$$

Problem 9.38. Let $\mathrm{sgn} : \mathbb{R} \to \mathbb{R}$ be the sign function defined by

$$\mathrm{sgn}(x) = \begin{cases} 1 & \text{if } x > 0\,, \\ -1 & \text{if } x < 0\,, \\ 0 & \text{if } x = 0\,. \end{cases}$$

Then clearly sgn is a tempered distribution since

$$\big|\langle \mathrm{sgn}, \phi \rangle\big| \leqslant \|\phi\|_{L^1(\mathbb{R})} \leqslant \pi p_2(\phi) \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,.$$

Show that $\dfrac{d}{dx}\mathrm{sgn}(x) = 2\delta$ in $\mathscr{S}(\mathbb{R})'$, where the derivative of tempered distributions is defined in Problem 9.36 and δ is the Dirac delta function.

Problem 9.39. 1. Let $a > 0$ be a constant. Find the Fourier transform of the function $f(x) = e^{-ax}H(x)$, where H is the Heaviside function defined by

$$H(x) = \mathbf{1}_{[0,\infty)}(x) = \begin{cases} 0 & \text{if } x < 0\,, \\ 1 & \text{if } x \geqslant 0\,. \end{cases}$$

2. Find the Fourier transform of the functions $g(x) = \dfrac{1}{2 - 3ix - x^2}$ by

(a) Rewriting g as the sum of two fractions and apply the result in part 1.

(b) Convolution.

Problem 9.40. Compute the Fourier transform of the function $f : \mathbb{R}^n \to \mathbb{R}$ given by $f(x) = |x|^\alpha$, where $-n < \alpha < 0$, by the following procedure.

1. Show that $f \notin L^1(\mathbb{R}^n)$.

2. Recall that the Gamma function $\Gamma : (0, \infty) \to \mathbb{R}$ defined by $\Gamma(x) = \displaystyle\int_0^\infty t^{x-1} e^{-t}\, dt$. Show that
$$|x|^\alpha = \frac{1}{\Gamma(-\frac{\alpha}{2})} \int_0^\infty s^{-\frac{\alpha}{2}-1} e^{-s|x|^2}\, ds \qquad \forall\, x \neq 0\,.$$

3. Find that Fourier transform of f.

4. Find the Fourier transform of the function $g : \mathbb{R}^n \to \mathbb{R}$ given by $g(x) = x_1|x|^\alpha$, where x_1 is the first component of x and $-n-2 < \alpha < -2$.

Hint: 3. Compute $\langle |x|^\alpha, \widehat{\phi}(x)\rangle$ by applying Fubini's Theorem several times.
4. Note that $g(x) = \dfrac{1}{\alpha+2} \dfrac{\partial}{\partial x_1} |x|^{\alpha+2}$ so that you can apply the results above. See Problem 9.36 for the Fourier transform of derivatives of tempered distributions.

Problem 9.41. Let $f \in L^1(\mathbb{R})$. Show that the function $y = \displaystyle\int_{-\infty}^x f(t)\, dt$ can be written as the convolution of f and a function $\phi \in L^1_{\text{loc}}(\mathbb{R})$.

Problem 9.42. In this problem we use symbolic computation to find the Fourier transform of the function
$$f(x) = \begin{cases} \dfrac{\sin(\omega x)}{x} & \text{if } x \neq 0\,, \\ \omega & \text{if } x = 0\,, \end{cases}$$
without knowing that it is the Fourier transform of the function $y = \sqrt{\dfrac{\pi}{2}} \mathbf{1}_{[-\omega,\omega]}(x)$. Complete the following.

1. Note that $f \notin L^1(\mathbb{R})$ but $f \in \mathscr{S}(\mathbb{R}^n)'$. Therefore, $\widehat{f} \in \mathscr{S}(\mathbb{R})$. Find the derivative of $\widehat{f}$, where the derivatives of tempered distributions is given in Problem 9.36.

2. Suppose that you can use the Fundamental Theorem of Calculus so that
$$\widehat{f}(\xi) - \widehat{f}(0) = \int_0^\xi \widehat{f}\,'(t)\,dt\,.$$
Note that in Problem 6.39 and 8.19 you are asked to show that $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$. Use this fact and treat $\delta_{\pm\omega}$ as the evaluation operation at $\pm\omega$ to find $\widehat{f}(\xi)$ (for $\xi \neq \pm\omega$).

Hint: 1. Recall that we have shown in Example 9.51 that $\mathscr{F}_x[\sin(\omega x)](\xi) = \dfrac{\sqrt{2\pi}}{2i}(\delta_\omega - \delta_{-\omega})$.

Problem 9.43. Let ω be a positive real number, and $f : \mathbb{R}^3 \to \mathbb{R}$ be defined by
$$f(x) = \begin{cases} \dfrac{\sin(\omega|x|)}{|x|} & \text{if } x \neq 0\,, \\ \omega & \text{if } x = 0\,, \end{cases}$$
where $|x| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ if $x = (x_1, x_2, x_3)$. In this problem we are concerned with the Fourier transform of f. Complete the following.

1. Show that $f \in \mathscr{S}(\mathbb{R}^3)'$.

2. Show that the Fourier transform of f is given by
$$\langle \widehat{f}, \varphi \rangle = \sqrt{\frac{\pi}{2}}\frac{1}{\omega}\int_{\partial B(0,\omega)} \varphi\,dS$$
for all $\varphi \in \mathscr{S}(\mathbb{R}^3)$, where $\displaystyle\int_{\partial B(0,\omega)} \varphi\,dS$ is the surface integral of φ on the sphere $\partial B(0,\omega)$ defined by
$$\int_{\partial B(0,\omega)} \varphi\,dS \equiv \int_0^\pi \int_0^{2\pi} \varphi(\omega\cos\theta\sin\phi, \omega\sin\theta\sin\phi, \omega\cos\phi)\omega^2 \sin\phi\,d\theta d\phi\,.$$

Hint of 2: You can show part 2 through the following procedures:

Step 1: By the definition of the Fourier transform of the tempered distributions,
$$\big\langle \widehat{f}, \varphi \big\rangle = \big\langle f, \widehat{\varphi} \big\rangle = \lim_{m\to\infty} \int_{B(0,m)} f(x)\Big(\frac{1}{\sqrt{2\pi}^3}\int_{\mathbb{R}^3} \varphi(\xi)e^{-ix\cdot\xi}\,d\xi\Big)\,dx$$
and the Fubini Theorem implies that
$$\big\langle \widehat{f}, \varphi \big\rangle = \frac{1}{\sqrt{2\pi}^3}\lim_{m\to\infty}\int_{\mathbb{R}^3}\Big(\int_{B(0,m)} f(x)e^{-ix\cdot\xi}\,dx\Big)\varphi(\xi)d\xi\,.$$
We focus on the inner integral first. Show that for each 3×3 orthonormal matrix O,
$$\int_{B(0,m)} f(x)e^{-ix\cdot\xi}\,dx = \int_{B(0,m)} \frac{\sin(\omega|y|)}{|y|}\,e^{-i(\mathrm{O}^{\mathrm{T}}\xi)\cdot y}\,dy\,.$$

Step 2: For each $\xi \in \mathbb{R}^3$, choose a 3×3 orthonormal matrix O such that $\mathrm{O}^{\mathrm{T}}\xi = (0, 0, |\xi|)$. Using the spherical coordinate $y = (\rho\cos\theta\sin\phi, \rho\sin\theta\sin\phi, \rho\cos\phi)$ to show that

$$\int_{B(0,m)} f(x)e^{-ix\cdot\xi}\,dx = \int_0^m \frac{2\sin(\omega\rho)\sin(|\xi|\rho)}{|\xi|}\,d\rho$$

so that we conclude that

$$\langle \widehat{f}, \varphi \rangle = \frac{1}{\sqrt{2\pi}^3} \lim_{m\to\infty} \int_{\mathbb{R}^3} \Big(\int_0^m \frac{2\sin(\omega\rho)\sin(|\xi|\rho)}{|\xi|} \varphi(\xi)\,d\rho \Big) d\xi\,.$$

Step 3: For each $r > 0$, define $\psi(r)$ as the surface integral of φ on $\partial B(0, r)$; that is,

$$\psi(r) = \int_{\partial B(0,r)} \varphi\,dS \equiv \int_0^\pi \int_0^{2\pi} \varphi(r\cos\theta\sin\phi, r\sin\theta\sin\phi, r\cos\phi) r^2 \sin\phi\,d\theta d\phi\,.$$

Using the spherical coordinate $\xi = (r\cos\theta\sin\phi, r\sin\theta\sin\phi, r\cos\phi)$ to show that

$$\langle \widehat{f}, \varphi \rangle = \frac{1}{\sqrt{2\pi}} \int_0^\infty \Big(\int_0^\infty \sin(\omega\rho)\sin(r\rho) \frac{2\psi(r)}{r}\,dr \Big) d\rho\,.$$

Step 4: Apply the conclusion in Problem 9.8.

Problem 9.44. 1. Show that the function $R : \mathbb{R} \to \mathbb{R}$ given by

$$\mathrm{R}(x) = \begin{cases} x & \text{if } x \geqslant 0\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

is a tempered distribution.

2. Let T be a generalized function defined by

$$\langle T, \phi \rangle = \lim_{\epsilon\to 0^+} \int_{\mathbb{R}\backslash[-\epsilon,\epsilon]} \frac{\phi(x)}{x}\,dx = \lim_{\epsilon\to 0^+} \Big(\int_{-\infty}^{-\epsilon} + \int_\epsilon^\infty \Big) \frac{\phi(x)}{x}\,dx \qquad \forall\,\phi \in \mathscr{C}_c^\infty(\mathbb{R})\,.$$

Show that $T \in \mathscr{S}(\mathbb{R})'$.

3. Let H be the Heaviside function given by

$$H(x) = \begin{cases} 0 & \text{if } x \leqslant 0\,, \\ 1 & \text{if } x > 0\,. \end{cases}$$

Show that $\widehat{H} = \dfrac{-i}{\sqrt{2\pi}} T + \sqrt{\dfrac{\pi}{2}}\,\delta$, here δ is the Dirac delta function.

Hint: 3. Let $G(x) = \exp\big(-\dfrac{x^2}{2}\big)$. For each $\phi \in \mathscr{S}(\mathbb{R})$, define $\psi = \phi - \phi(0)G$ (which belongs to $\mathscr{S}(\mathbb{R})$). Use the identity

$$\langle \widehat{H}, \phi\rangle = \langle H, \widehat{\psi}\rangle - \phi(0)\langle H, \widehat{G}\rangle$$

to make the conclusion.

Problem 9.45. The Hilbert transform of a function $f : \mathbb{R} \to \mathbb{R}$, denoted by $\mathscr{H}[f]$, is a function defined (formally) by

$$\mathscr{H}[f](x) = \frac{1}{\pi} \lim_{\epsilon \to 0^+} \int_{|y-x|>\epsilon} \frac{f(y)}{x-y}\, dy\,,$$

1. Show that $\mathscr{H}[f]$ is well-defined if $f \in \mathscr{S}(\mathbb{R})$.
2. Show that $\mathscr{F}\big[\mathscr{H}[f]\big](\xi) = i\mathrm{sgn}(\xi)\widehat{f}(\xi)$ for all $f \in \mathscr{S}(\mathbb{R})$.
3. Show that $\big\|\mathscr{H}[f]\big\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})}$ for all $f \in \mathscr{S}(\mathbb{R})$, where $\|g\|_{L^2(\mathbb{R})} = \Big(\displaystyle\int_{\mathbb{R}} \big|g(x)\big|^2\, dx\Big)^{\frac{1}{2}}$.

Hint: Consider the tempered distribution T defined in Problem 9.44 by

$$\langle T, \phi\rangle = \lim_{\epsilon \to 0^+} \int_{\mathbb{R}\backslash[-\epsilon,\epsilon]} \frac{\phi(x)}{x}\, dx = \lim_{\epsilon \to 0^+} \Big(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty}\Big) \frac{\phi(x)}{x}\, dx \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,.$$

1. Show that $\mathscr{H}[f] = \langle T, \tau_x \widetilde{f}\rangle$ for all $f \in \mathscr{S}(\mathbb{R})$, where τ_x is a translation operator.
2. Show that the tempered distribution S defined by $\langle S, \phi\rangle = \langle T(x), x\phi(x)\rangle$ is indeed the same as the tempered distribution

   $$\phi \mapsto \int_{\mathbb{R}} \phi(x)\, dx = \langle 1, \phi\rangle\,.$$

   Use Problem 9.36 to show that $\dfrac{d}{d\xi}\widehat{T}(\xi) = -\sqrt{\dfrac{\pi}{2}}\, i \dfrac{d}{d\xi}\mathrm{sgn}(\xi)$, where sgn is given in Problem 9.38. Use the fact that $\dfrac{dT}{dx} = 0$ if and only if there exists C such that $\langle T, \phi\rangle = \langle C, \phi\rangle$ for all $\phi \in \mathscr{S}(\mathbb{R})$ to conclude that

   $$\widehat{T}(\xi) = -\sqrt{\frac{\pi}{2}}\, i\mathrm{sgn}(\xi) + C$$

   for some constant C. Find the constant C and also show that $\mathscr{H}[f] = \dfrac{1}{\pi} T * f = \sqrt{\dfrac{2}{\pi}}\, T \star f$.

3. Use the Plancherel formula.

Problem 9.46 (Fourier-analytic direction for generalized ridge density)**.** In this problem you are asked to prove the "$\Rightarrow$" direction of Proposition 7.90:

Proposition 7.90 *If $\mathcal{M}(\Omega)$ is dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$ in the sense of uniform convergence on compact subsets, then the only polynomial in H^n that vanishes identically on $L(\Omega)$ is the zero polynomial.*

Use the notation and setting from Problem 7.23. This completes the proof of Proposition 7.67.

Hint. Argue by contradiction. Suppose there exist $k \in \mathbb{N}$ and $p \in H_k^n \backslash \{0\}$ such that $p(\xi) = 0$ for all $\xi \in L(\Omega)$. Write

$$p(\xi) = \sum_{|\boldsymbol{m}|=k} b_{\boldsymbol{m}}\, \xi^{\boldsymbol{m}}.$$

Choose a nontrivial $\phi \in \mathscr{C}_c^\infty(\mathbb{R}^n;\mathbb{R})$ and define

$$\psi(\boldsymbol{x}) := \sum_{|\boldsymbol{m}|=k} b_{\boldsymbol{m}}\, D^{\boldsymbol{m}}\phi(\boldsymbol{x}).$$

Then $\psi \in \mathscr{C}_c^\infty(\mathbb{R}^n;\mathbb{R})$, $\psi \neq 0$, and

$$\widehat{\psi}(\xi) = i^k\, \widehat{\phi}(\xi)\, p(\xi).$$

Show that ψ defines a nontrivial continuous linear functional on $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$ that annihilates $\mathcal{M}(\Omega)$.

1. Fix $A \in \Omega$ and let $m := \dim L(A)$. After an orthogonal change of coordinates, write $\boldsymbol{x} = (\boldsymbol{x}', \boldsymbol{x}'')$ with $\boldsymbol{x}' \in \mathbb{R}^m$ and $\boldsymbol{x}'' \in \mathbb{R}^{n-m}$ so that $L(A) = \mathbb{R}^m \times \{\boldsymbol{0}\}$. Define

$$H(\boldsymbol{x}') := \int_{\mathbb{R}^{n-m}} \psi(\boldsymbol{x}', \boldsymbol{x}'')\, d\boldsymbol{x}''.$$

   Show that $H \equiv 0$ by proving $\widehat{H}(\boldsymbol{c}') = 0$ for all $\boldsymbol{c}' \in \mathbb{R}^m$.

2. Use the fact that $A(\boldsymbol{x}', \boldsymbol{x}'') = A(\boldsymbol{x}', 0)$ to show that for any $g \in \mathscr{C}(\mathbb{R}^d;\mathbb{R})$,

$$\int_{\mathbb{R}^n} g(A\boldsymbol{x})\, \psi(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\mathbb{R}^m} H(\boldsymbol{x}')\, g(A(\boldsymbol{x}', 0))\, d\boldsymbol{x}'.$$

   Conclude that

$$\int_{\mathbb{R}^n} g(A\boldsymbol{x})\, \psi(\boldsymbol{x})\, d\boldsymbol{x} = 0 \qquad \forall\, A \in \Omega,\ \forall\, g \in \mathscr{C}(\mathbb{R}^d;\mathbb{R}). \tag{9.5.3}$$

3. Deduce from (9.5.3) that $\mathcal{M}(\Omega)$ cannot be dense in $\mathscr{C}(\mathbb{R}^n;\mathbb{R})$.

Problem 9.47 (A Baire-category proof of UAT)**.** In this problem we establish the "$\Rightarrow$" direction of the Universal Approximation Theorem by combining Theorem 7.91 with the Baire Category Theorem (Theorem 7.82).

Let $\sigma \in \mathscr{C}(\mathbb{R};\mathbb{R})$ and assume that σ is not a polynomial. Denote by

$$\mathscr{C}_c^\infty(\mathbb{R};\mathbb{R}) = \left\{\eta : \mathbb{R} \to \mathbb{R} \,\middle|\, \eta \in \mathscr{C}^\infty(\mathbb{R};\mathbb{R}),\ \operatorname{supp}(\eta) := \overline{\{x \in \mathbb{R} \mid \eta(x) \neq 0\}} \text{ is compact}\right\}$$

the space of smooth compactly supported functions. Complete the following steps.

1. Show that for every $\eta \in \mathscr{C}_c^\infty(\mathbb{R};\mathbb{R})$, the convolution

$$(\sigma * \eta)(x) := \int_{\mathbb{R}} \sigma(y)\, \eta(x-y)\, dy$$

belongs to $\mathscr{C}^\infty(\mathbb{R};\mathbb{R}) \cap \overline{\Sigma_1(\sigma)}^{\|\cdot\|_\infty}$, where the closure is taken with respect to uniform convergence on compact subsets.

Hint. For smoothness, differentiate under the integral sign and apply the Dominated Convergence Theorem (Theorem 6.102) or the Bounded Convergence Theorem (Corollary 6.68). For density, approximate the integral defining $\sigma * \eta$ by Riemann sums and observe that each Riemann sum is a finite linear combination of translates of σ.

2. Assume that $\sigma * \eta$ is a polynomial for every $\eta \in \mathscr{C}_c^\infty(\mathbb{R};\mathbb{R})$. Show that there exists $m \in \mathbb{Z}_+$ such that $\sigma * \eta$ is a polynomial of degree at most m for all $\eta \in \mathscr{C}_c^\infty(\mathbb{R};\mathbb{R})$.

Hint. For each $k \in \mathbb{Z}_+$, define

$$V_k = \left\{\eta \in \mathscr{C}_c^\infty(\mathbb{R};\mathbb{R}) \,\middle|\, \sigma * \eta \text{ is a polynomial of degree at most } k\right\}.$$

Then

$$\bigcup_{k=0}^{\infty} V_k = \mathscr{C}_c^\infty(\mathbb{R};\mathbb{R}).$$

Use the Baire Category Theorem to show that some V_k has nonempty interior. Since each V_k is a vector space, deduce that this V_k must coincide with the whole space. (To make this argument rigorous, equip $\mathscr{C}_c^\infty(\mathbb{R};\mathbb{R})$ with a suitable complete metric; for example, use the standard countable family of seminorms on compact sets.)

3. Show that if there exists $m \in \mathbb{Z}_+$ such that $\sigma * \eta$ is a polynomial of degree at most m for all $\eta \in \mathscr{C}_c^\infty(\mathbb{R};\mathbb{R})$, then σ itself is a polynomial of degree at most m.

   Hint. Use a standard mollifier (Definition 9.30) and apply Lemma 9.34 to pass to the limit as the mollifier concentrates at a point.

4. Combine the previous steps with Theorem 7.91 to complete the proof of the "$\Rightarrow$" direction of the Universal Approximation Theorem.

§ 9.5 Chapter Review

Problem 9.48 (**True or False**). Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. If the Cesàro mean $\{b_n\}_{n=1}^\infty$ of a sequence $\{a_k\}_{k=1}^\infty$ converges, the sequence $\{a_k\}_{k=1}^\infty$ also converges.

Chapter 10

Applications

10.1 Application on Signal Processing - the Sampling Theorem and the Nyquist Rate

In the study of signal processing, the Fourier transform and the inverse Fourier transform are often defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{-2\pi i x\cdot\xi}dx \quad \text{and} \quad \check{f}(x) = \int_{\mathbb{R}^n} f(\xi)e^{2\pi i x\cdot\xi}d\xi \qquad \forall\, f \in L^1(\mathbb{R}^n). \tag{10.1.1}$$

For $T \in \mathscr{S}(\mathbb{R}^n)'$, the Fourier transform of T is defined again by

$$\langle \widehat{T}, \phi\rangle = \langle T, \widehat{\phi}\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).$$

We also note that the definitions of the translation, dilation, and reflection of tempered distributions are independent of the Fourier transform, and are still defined by

$$\langle \tau_h T, \phi\rangle = \langle T, \tau_{-h}\phi\rangle, \quad \langle d_\lambda T, \phi\rangle = \langle T, \lambda^n d_{\lambda^{-1}}\phi\rangle \quad \text{and} \quad \langle \widetilde{T}, \phi\rangle = \langle T, \widetilde{\phi}\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).$$

Concerning the convolution, when the Fourier transform is given by (10.1.1), we usually consider the $\circledast$ convolution operator

$$(f \circledast g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)\,dy = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy \qquad \forall\, f, g \in L^1(\mathbb{R}^n).$$

instead of $\star$ convolution operators. The convolution of T and $f \in \mathscr{S}(\mathbb{R}^n)$ is defined by

$$\langle T \circledast f, \phi\rangle = \langle T, \widetilde{f} \circledast \phi\rangle = \langle \widetilde{T}, f \circledast \widetilde{\phi}\rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R}^n).$$

Then similar to Theorem 9.52, 9.54, and 9.59, we have

1. $\check{\hat{T}} = \hat{\check{T}} = T$ for all $T \in \mathscr{S}(\mathbb{R}^n)'$.

2. $\widehat{\tau_h T}(\xi) = \widehat{T}(\xi)e^{-2\pi i\xi\cdot h}$, $\widehat{d_\lambda T} = \lambda^n d_{\frac{1}{\lambda}}\widehat{T}$, and $\hat{\hat{T}} = \check{T}$ for all $T \in \mathscr{S}(\mathbb{R}^n)'$.

3. $\widehat{T * f} = \widehat{T}\widehat{f}$ and $\widehat{fT} = \widehat{f} * \widehat{T}$ for all $f \in \mathscr{S}(\mathbb{R}^n)$ and $T \in \mathscr{S}(\mathbb{R}^n)'$. Moreover, if $S \in \mathscr{S}(\mathbb{R}^n)'$ has the property that $S * \phi \in \mathscr{S}(\mathbb{R}^n)$ for all $\phi \in \mathbb{R}^n$, then $\widehat{T * S} = \widehat{T}\widehat{S}$ in $\mathscr{S}(\mathbb{R}^n)'$ for all $T \in \mathscr{S}(\mathbb{R}^n)'$.

Moreover,

1. $\widehat{\delta} = \check{\delta} = 1$ in $\mathscr{S}(\mathbb{R}^n)'$, and $\widehat{\delta_h}(\xi) = \widehat{\tau_h\delta}(\xi) = \widecheck{\delta_{-h}} = \widecheck{\tau_{-h}\delta} = e^{-2\pi i h\cdot\xi}$ in $\mathscr{S}(\mathbb{R}^n)'$ for all $h \in \mathbb{R}^n$.

2. By Euler's identity, $\widehat{\cos(2\pi\omega x)}(\xi) = \dfrac{1}{2}(\delta_\omega + \delta_{-\omega})$ and $\widehat{\sin(2\pi\omega x)}(\xi) = \dfrac{1}{2i}(\delta_\omega - \delta_{-\omega})$.

3. $\delta * \delta = \delta$, and $\delta_a * \delta_b = \delta_{a+b}$ for all $a, b \in \mathbb{R}^n$.

4. $\delta * \phi = \phi$ and $\delta_a * \phi = \tau_a\phi$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$.

5. Re-define the rect function $\Pi : \mathbb{R} \to \mathbb{R}$ by
$$\Pi(x) = \begin{cases} 1 & \text{if } |x| < \dfrac{1}{2}, \\ 0 & \text{if } |x| \geqslant \dfrac{1}{2}. \end{cases} \tag{10.1.2}$$
Then $\widehat{\Pi}(\xi) = \check{\Pi}(\xi) = \operatorname{sinc}(\xi)$, where sinc is the normalized sinc function given by (9.4.1).

6. Let $\Lambda : \mathbb{R} \to \mathbb{R}$ be the triangle function define by
$$\Lambda(x) = \begin{cases} 1 - |x| & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geqslant 1. \end{cases}$$
Then by the fact that Λ is an even function, if $\xi \neq 0$,
$$\begin{aligned} \widehat{\Lambda}(\xi) &= 2\int_0^1 (1-x)\cos(2\pi x\xi)\,dx = 2\Big[(1-x)\frac{\sin(2\pi x\xi)}{2\pi\xi}\Big|_{x=0}^{x=1} + \int_0^1 \frac{\sin(2\pi x\xi)}{2\pi\xi}\,dx\Big] \\ &= \frac{1-\cos(2\pi\xi)}{2\pi^2\xi^2} = \frac{\sin^2\pi\xi}{\pi^2\xi^2}, \end{aligned}$$
while $\widehat{\Lambda}(0) = 1$. Therefore, $\widehat{\Lambda}(\xi) = \operatorname{sinc}^2(\xi)$. Using the property of convolution, we have $\Pi * \Pi = \Lambda$.

When a continuous function, $x(t)$, is sampled at a constant rate F_s samples per second（以每秒 F_s 次均匀取樣）, there is always an unlimited number of other continuous functions that fit the same set of samples; however, only one of them is bandlimited to $\frac{1}{2}F_s$ cycles per second (hertz), which means that its Fourier transform, $\widehat{x}(f)$, is 0 for all $|f| \geqslant \frac{1}{2}F_s$.

Definition 10.1. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. f is said to be a ***bandlimited*** function if $\operatorname{supp}(\widehat{f})$ is bounded. The ***bandwidth*** of a bandlimited function f is the number $\sup \operatorname{supp}(\widehat{f})$. f is said to be ***timelimited*** if $\operatorname{supp}(f)$ is bounded.

Definition 10.2. In signal processing, the ***Nyquist rate*** is twice the bandwidth of a bandlimited function or a bandlimited channel.

In the field of digital signal processing, the ***sampling theorem*** is a fundamental bridge between continuous-time signals (often called "analog signals") and discrete-time signals (often called "digital signals"). It establishes a sufficient condition for a *sample rate*（取樣頻率）that permits a discrete sequence of samples to capture all the information from a continuous-time signal of finite bandwidth. To be more precise, Shannon's version of the theorem states that "if an analog signal contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced $\dfrac{1}{2B}$ seconds apart."

In the following, we examine the sampling theorem rigorously. We start with the simplest version that the signal is continuous, integrable and bandlimited.

Theorem 10.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous integrable function. If $\operatorname{supp}(\widehat{f}) \subseteq [-B, B]$, then f is fully determined by the sequence $\Big\{ f\big(\frac{k}{2B}\big) \Big\}_{k=-\infty}^{\infty}$, and*

$$f(x) = \sum_{k=-\infty}^{\infty} f\Big(\frac{k}{2B}\Big) \operatorname{sinc}(2Bx - k) \qquad \forall\, x \in \mathbb{R}. \tag{10.1.3}$$

Proof. Let $f \in L^1(\mathbb{R}) \cap \mathscr{C}(\mathbb{R};\mathbb{R})$ such that $\operatorname{supp}(\widehat{f}) \subseteq [-B, B]$. Then $\widehat{f} \in L^1(\mathbb{R})$ (since it must be bounded by Proposition 9.4); thus the Fourier inversion formula implies that

$$f(x) = \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i x \xi}\, d\xi = \int_{-B}^{B} \widehat{f}(\xi) e^{2\pi i x \xi}\, d\xi \qquad \forall\, x \in \mathbb{R}.$$

In particular,

$$f\Big(\frac{k}{2B}\Big) = \int_{-B}^{B} \widehat{f}(\xi) e^{\frac{i\pi k \xi}{B}}\, d\xi.$$

Treating $\widehat{f}$ as a function defined on $[-B,B]$, the identity above implies that $\Big\{\frac{1}{2B}f\big(\frac{-k}{2B}\big)\Big\}_{k=-\infty}^{\infty}$ is the Fourier coefficients of $\widehat{f}$.

Note that the boundedness of $\widehat{f}$ implies that $\widehat{f}\in L^2(-B,B)$. Therefore, if $g\in L^2(-B,B)$, the Parseval identity (or (8.5.5)) implies that

$$\frac{1}{2B}\int_{-B}^{B}\widehat{f}(\xi)\overline{g(\xi)}\,d\xi=\langle\widehat{f},g\rangle_{L^2(-B,B)}=\sum_{k=-\infty}^{\infty}\frac{1}{2B}f\big(\frac{-k}{2B}\big)\,\overline{\widehat{g}_k}\,,\tag{10.1.4}$$

where $\widehat{g}_k=\dfrac{1}{2B}\displaystyle\int_{-B}^{B}g(x)e^{-\frac{i\pi kx}{B}}\,dx$.

For each $x\in\mathbb{R}$, the Fourier coefficients of the function $g(\xi)=e^{-2\pi ix\xi}$ is given by

$$\widehat{g}_k=\frac{1}{2B}\int_{-B}^{B}e^{-2\pi ix\xi}e^{-\frac{i\pi k\xi}{B}}\,d\xi=\frac{1}{2B}\int_{-B}^{B}e^{-i\frac{\pi(2Bx+k)}{B}\xi}\,d\xi=\operatorname{sinc}(2Bx+k)\,;$$

thus the Fourier inversion formula and (10.1.4) imply that if $x\in\mathbb{R}$,

$$f(x)=\int_{-B}^{B}\widehat{f}(\xi)\overline{g(\xi)}\,dx=\sum_{k=-\infty}^{\infty}f\big(\frac{-k}{2B}\big)\operatorname{sinc}(2Bx+k)=\sum_{k=-\infty}^{\infty}f\big(\frac{k}{2B}\big)\operatorname{sinc}(2Bx-k)\,.$$

The identity above shows that f is fully determined by the sequence $\Big\{f\big(\frac{k}{2B}\big)\Big\}_{k=-\infty}^{\infty}$. $\square$

Remark 10.4. Equation (10.1.3) is called the ***Whittaker–Shannon interpolation formula***.

Remark 10.5. Suppose that $f\in\mathscr{C}(\mathbb{R};\mathbb{R})\cap\mathscr{S}(\mathbb{R})'$ be such that $\widehat{f}$, in the sense of tempered distribution, belongs to $L^2(\mathbb{R})$ and is supported in $[-B,B]$ $\big($that is, $\big\{x\in\mathbb{R}\,\big|\,\widehat{f}(x)\neq 0\big\}\subseteq[-B,B]\big)$. By the definition of the Fourier transform for $\mathscr{S}(\mathbb{R})'$ we have

$$\big\langle\check{\widehat{f}}-f,\phi\big\rangle=0\qquad\forall\,\phi\in\mathscr{S}(\mathbb{R});$$

thus by the fact that $\check{\widehat{f}}\in\mathscr{C}_b(\mathbb{R};\mathbb{R})$, Lemma 9.35 implies that

$$f(x)=\check{\widehat{f}}(x)=\int_{-B}^{B}\widehat{f}(\xi)e^{2\pi i\xi x}\,d\xi\qquad\forall\,x\in\mathbb{R}\,.$$

Therefore, the Fourier coefficients of $\widehat{f}$ is again the sequence $\Big\{\frac{1}{2B}f\big(\frac{-k}{2B}\big)\Big\}_{k=-\infty}^{\infty}$ so that the same argument of showing Theorem 10.3 establishes (10.1.3).

Remark 10.6 (The inner product point of view). Let $e_k(x) = \text{sinc}(x-k) = (\tau_k\text{sinc})(x)$. Then $e_k \in L^2(\mathbb{R})$ since

$$\int_{\mathbb{R}} \text{sinc}^2(x-k)\,dx = \int_{\mathbb{R}} \text{sinc}^2 x dx = \int_{\mathbb{R}} \frac{\sin^2 \pi x}{\pi^2 x^2} dx = \frac{1}{\pi}\int_{\mathbb{R}} \frac{\sin^2 x}{x^2} dx < \infty\,.$$

By the Plancherel formula (9.3.5),

$$\langle e_k, e_\ell\rangle_{L^2(\mathbb{R})} = \big\langle \widetilde{\tau_k\text{sinc}}, \widetilde{\tau_\ell\text{sinc}}\big\rangle_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \Pi(\xi)e^{2\pi ik\xi}\overline{\Pi(\xi)e^{2\pi i\ell\xi}}\,d\xi = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i(k-\ell)\xi}\,d\xi$$

which is 0 if $k \neq \ell$ and is 1 is $k = \ell$. Therefore, we find that $\{e_k\}_{k\in\mathbb{Z}}$ is an orthonormal set in $L^2(\mathbb{R})$.

Now suppose that $f \in L^2(\mathbb{R})$ (so that $\widehat{f} \in L^2(\mathbb{R})$ by the Plancherel formula) such that $\text{supp}(\widehat{f}) \subseteq \big(-\frac{1}{2}, \frac{1}{2}\big)$. Then

$$\begin{aligned}\langle f, e_k\rangle_{L^2(\mathbb{R})} &= \big\langle \widehat{f}, \widehat{\tau_k \sin c}\big\rangle_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \widehat{f}(\xi)\overline{\Pi(-\xi)e^{-2\pi ik\xi}}\,d\xi = \int_{-\frac{1}{2}}^{\frac{1}{2}} \widehat{f}(\xi)e^{2\pi ik\xi}\,d\xi \\ &= \int_{\mathbb{R}} \widehat{f}(\xi)e^{2\pi ik\xi}\,d\xi = \check{\widehat{f}}(k) = f(k)\end{aligned}$$

if f is continuous at k. By Remark 10.5, if $f \in L^2(\mathbb{R}) \cap \mathscr{C}(\mathbb{R})$ such that $\text{supp}(\widehat{f}) \subseteq \big(-\frac{1}{2}, \frac{1}{2}\big)$, then

$$f(x) = \sum_{k=-\infty}^{\infty} f(k)\text{sinc}(x-k) = \sum_{k=-\infty}^{\infty} (f, e_k)_{L^2(\mathbb{R})}e_k(x) \qquad \forall\, x \in \mathbb{R}.$$

In other words, one can treat $\{e_k\}_{k\in\mathbb{Z}}$ as an "orthonormal basis" in the space

$$\Big\{ f \in L^2(\mathbb{R})\big(\cap\mathscr{C}(\mathbb{R})\big)\,\Big|\,\text{supp}(\widehat{f}) \subseteq \big(-\frac{1}{2}, \frac{1}{2}\big)\Big\}.$$

10.2 Application on Partial Differential Equations

10.2.1 Heat Conduction in a Rod

Consider the heat distribution on a rod of length L: Parameterize the rod by $[0, L]$, and let t be the time variable. Let $\rho(x)$, $s(x)$, $\kappa(x)$ denote the density, specific heat, and the thermal conductivity of the rod at position $x \in (0, L)$, respectively, and $u(x,t)$ denote the temperature at position x and time t. For $0 < x < L$, and $\Delta x, \Delta t \ll 1$,

$$\int_x^{x+\Delta x} \rho(y)s(y)\big[u(y,t+\Delta t)-u(y,t)\big]\,dy = \int_t^{t+\Delta t} \big[-\kappa(x)u_x(x,t')+\kappa(x+\Delta x)u_x(x+\Delta x,t')\big]\,dt',$$

where the left-hand side denotes the change of the total heat in the small section $(x, x+\Delta x)$, and the right-hand side denotes the heat flows from outside. Divide both sides by $\Delta x \Delta t$ and letting Δx and Δt approach zero, if all the functions appearing in the equation above are smooth enough, we find that

$$\rho(x)s(x)u_t(x,t) = \big[\kappa(x)u_x(x,t)\big]_x \qquad 0 < x < L, \quad t > 0. \tag{10.2.1}$$

Assuming uniform rod; that is, ρ, s, κ are constant, then (10.2.1) reduces to that

$$u_t(x,t) = \alpha^2 u_{xx}(x,t), \qquad 0 < x < L, \quad t > 0, \tag{10.2.2a}$$

where $\alpha^2 = \dfrac{\kappa}{\rho s}$ is called the ***thermal diffusivity***. By re-scaling the length scale, we can assume that $\alpha = 1$.

To determine the state of the temperature, we need to impose that initial condition

$$u(x,0) = u_0(x) \qquad 0 < x < L \tag{10.2.2b}$$

for some given function $u_0 : [0, L] \to \mathbb{R}$ and a boundary condition. Usually one of the following four types of boundary conditions is imposed:

1. ***Dirichlet boundary condition***: The Dirichlet boundary condition is used to describe the phenomena that the temperature at the end points of the rod is known/ controable. Mathematically, it is expressed by

   $$u(0,t) = a(t) \quad \text{and} \quad u(L,t) = b(t) \qquad \forall\, t > 0$$

   for some given functions $a(t)$ and $b(t)$.

2. ***Neumann boundary condition***: The Neumann boundary condition is used to describe the phenomena of insulation; that is, there is no heat flow at the end points. Mathematically, it is expressed by

   $$u_x(0,t) = u_x(L,t) = 0 \qquad \forall\, t > 0.$$

   In general, we can consider the boundary condition

   $$u_x(0,t) = a(t) \quad \text{and} \quad u_x(L,t) = b(t) \qquad \forall\, t > 0$$

   for some given functions $a(t)$ and $b(t)$.

3. ***Mixed type boundary condition***: We can also consider the case that at one end point the temperature is known while there is no heat flow on the other end point. In general, this is expressed by

$$u(0,t) = a(t) \quad \text{and} \quad u_x(L,t) = b(t) \qquad \forall\, t > 0$$

or

$$u_x(0,t) = a(t) \quad \text{and} \quad u(L,t) = b(t) \qquad \forall\, t > 0$$

for some given functions $a(t)$ and $b(t)$.

4. ***Periodic boundary condition***: Suppose that instead of rods we consider modelling the temperature distribution in a (big) ring (with perimeter L). Choosing a point on the ring as the "left-end" point and parameterizing the point of the ring by arc-length, we then have the "boundary" condition

$$u(0,t) = u(L,t) \qquad \forall\, t > 0\,.$$

This is called the periodic boundary condition.

The Dirichlet problem

In this sub-section we consider the heat equation with Dirichlet boundary condition:

$$\begin{aligned} u_t - u_{xx} &= 0 && \text{in} \quad (0,L)\times(0,\infty)\,, && \text{(10.2.3a)}\\ u &= u_0 && \text{on} \quad (0,L)\times\{t=0\}\,, && \text{(10.2.3b)}\\ u(0,t) = a\,, \quad u(L,t) &= b && \text{for all } t>0\,, && \text{(10.2.3c)} \end{aligned}$$

where a and b are given constants. Let $v(x,t) = u(x,t) - \dfrac{b-a}{L}x - a$. Then v satisfies

$$\begin{aligned} v_t - v_{xx} &= 0 && \text{in} \quad (0,L)\times(0,\infty)\,, && \text{(10.2.4a)}\\ v &= v_0 && \text{on} \quad (0,L)\times\{t=0\}\,, && \text{(10.2.4b)}\\ v(0,t) = v(L,t) &= 0 && \text{for all } t>0\,, && \text{(10.2.4c)} \end{aligned}$$

where $v_0 : [0,L] \to \mathbb{R}$ is given by $v_0(x) = u_0(x) - \dfrac{b-a}{L}x - a$. As long as the solution v to (10.2.4) is found, the solution u to (10.2.3) can be constructed using $u(x,t) = v(x,t) + \dfrac{b-a}{L}x + a$. Therefore, we focus on solving (10.2.4) (using the Fourier series method).

The idea of using the Fourier series to solve (10.2.4) is that for each fixed $t > 0$ we express v in terms of its Fourier series representation (using proper "basis"). Recall that for a function $f : [0, L] \to \mathbb{R}$, we have the Fourier representation

$$f(x) \text{``}=\text{''} \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k \cos \frac{2\pi kx}{L} + s_k \sin \frac{2\pi kx}{L} \qquad x \in [0, L],$$

where $c_k = \dfrac{2}{L}\displaystyle\int_0^L f(x) \cos \frac{2\pi kx}{L}\,dx$ and $s_k = \dfrac{2}{L}\displaystyle\int_0^L f(x) \sin \frac{2\pi kx}{L}\,dx$, so

$$v(x,t) \text{``}=\text{''} \frac{c_0(t)}{2} + \sum_{k=1}^{\infty} c_k(t) \cos \frac{2\pi kx}{L} + s_k(t) \sin \frac{2\pi kx}{L} \qquad x \in [0, L],$$

for some sequence of functions $\big\{c_k(t)\big\}_{k=0}^{\infty}$ and $\big\{s_k(t)\big\}_{k=1}^{\infty}$. However, this particular Fourier series representation of v is not a good choice of solving (10.2.4) since it is difficult to validate the boundary condition (10.2.4c).

Note that for $f : [0, L] \to \mathbb{R}$ instead of the Fourier series representation above we can also consider the "cosine" series or "sine" series that are obtained by treating f as the restriction of an even or an odd function defined on $[-L, L]$ to $[0, L]$. In other words, define $f_e, f_o : [-L, L] \to \mathbb{R}$, called the ***even*** and ***odd extension*** of f respectively, by

$$f_e(x) = \begin{cases} f(x) & \text{if } x \in [0, L], \\ f(-x) & \text{if } x \in [-L, 0), \end{cases} \qquad \text{and} \qquad f_o(x) = \begin{cases} f(x) & \text{if } x \in [0, L], \\ -f(-x) & \text{if } x \in [-L, 0), \end{cases}$$

then $f = f_o = f_e$ on $[0, L]$. Since

$$f_e(x) \text{``}=\text{''} \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k \cos \frac{\pi kx}{L} \qquad \text{and} \qquad f_o(x) \text{``}=\text{''} \sum_{k=1}^{\infty} s_k \sin \frac{\pi kx}{L} \qquad x \in [-L, L],$$

where $c_k = \dfrac{2}{L}\displaystyle\int_0^L f(x) \cos \frac{\pi kx}{L}\,dx$ and $s_k = \dfrac{2}{L}\displaystyle\int_0^L f(x) \sin \frac{\pi kx}{L}\,dx$, we have

$$f(x) \text{``}=\text{''} \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k \cos \frac{\pi kx}{L} \qquad \text{and} \qquad f(x) \text{``}=\text{''} \sum_{k=1}^{\infty} s_k \sin \frac{\pi kx}{L} \qquad x \in [0, L].$$

Using the sine series, for each $t > 0$ $v(x,t)$ can be expressed as

$$v(x,t) \text{``}=\text{''} \sum_{k=1}^{\infty} d_k(t) \sin \frac{\pi kx}{L} \qquad x \in [0, L].$$

for some sequence of function $\{d_k(t)\}_{k=1}^{\infty}$ to be determined. We note that using this particular representation of v the boundary condition (10.2.4c) automatically holds. Therefore, it suffices to find $\{d_k(t)\}_{k=1}^{\infty}$ such that (10.2.4a,b) hold.

Assume that the differentiation of the series can be obtained by term-by-term differentiation; that is,

$$\frac{\partial}{\partial t}\sum_{k=1}^{\infty} d_k(t)\sin\frac{\pi kx}{L} = \sum_{k=1}^{\infty}\frac{\partial}{\partial t}\big(d_k(t)\sin\frac{\pi kx}{L}\big) = \sum_{k=1}^{\infty} d_k'(t)\sin\frac{\pi kx}{L}$$

and

$$\frac{\partial^2}{\partial x^2}\sum_{k=1}^{\infty} d_k(t)\sin\frac{\pi kx}{L} = \sum_{k=1}^{\infty}\frac{\partial^2}{\partial x^2}\big(d_k(t)\sin\frac{\pi kx}{L}\big) = -\sum_{k=1}^{\infty}\frac{k^2\pi^2}{L^2}d_k(t)\sin\frac{\pi kx}{L}.$$

Then (10.2.4a) implies that

$$\sum_{k=1}^{\infty}\Big[d_k'(t) + \frac{k^2\pi^2}{L^2}d_k(t)\Big]\sin\frac{\pi kx}{L} = 0 \qquad x \in [0, L].$$

As a consequence,

$$d_k'(t) + \frac{k^2\pi^2}{L^2}d_k(t) = 0 \qquad \forall\, k \in \mathbb{N}. \tag{10.2.5a}$$

To determine d_k uniquely, an initial condition for d_k has to be imposed. Noting that (10.2.4b) implies that

$$v_0(x)\,\text{``}=\text{''}\sum_{k=1}^{\infty} d_k(0)\sin\frac{\pi kx}{L} \qquad x \in [0, L];$$

thus

$$d_k(0) = \widehat{v}_{0k} \equiv \frac{2}{L}\int_0^L v_0(x)\sin\frac{\pi kx}{L}\,dx. \tag{10.2.5b}$$

Solving the initial value problem (10.2.5), we find that

$$d_k(t) = \widehat{v}_{0k}e^{-\frac{k^2\pi^2}{L^2}t} \qquad \forall\, k \in \mathbb{N};$$

thus the solution to (10.2.4) can be written as

$$v(x, t) = \sum_{k=1}^{\infty}\widehat{v}_{0k}e^{-\frac{k^2\pi^2}{L^2}t}\sin\frac{k\pi x}{L}.$$

Therefore, the solution to (10.2.3) can be written as

$$u(x, t) = \sum_{k=1}^{\infty}\widehat{v}_{0k}e^{-\frac{k^2\pi^2}{L^2}t}\sin\frac{k\pi x}{L} + \frac{b-a}{L}x + a. \tag{10.2.6}$$

• the long time behavior: Suppose that the temperature at the left-end and right-end points are fixed as a and b (as described in the boundary condition (10.2.3c)). Then we expect that no matter what the temperature distribution is given initially, the temperature distribution approaches a linear distribution; that is, we expect that $u(x,t) \to \dfrac{b-a}{L}x + a$ as $t \to \infty$ for all $x \in [0,L]$. This expectation is in fact true, and we try to prove this here.

Using (10.2.6), we obtain that

$$\Big|u(x,t) - \frac{b-a}{L}x - a\Big| \leqslant \sum_{k=1}^{\infty} \big|\widehat{v}_{0k}\big| e^{-\frac{k^2\pi^2}{L^2}t}.$$

By the fact that

$$\big|\widehat{v}_{0k}\big| \leqslant \frac{2}{L}\int_0^L \big|v_0(x)\big|\,dx = \frac{2}{L}\|v_0\|_{L^1(0,L)},$$

we find that

$$\begin{aligned}\Big|u(x,t) - \frac{b-a}{L}x - a\Big| &\leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)}\sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}t} \leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)}\sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}(t-1)}e^{-\frac{k^2\pi^2}{L^2}} \\ &\leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)} e^{-\frac{\pi^2}{L^2}(t-1)}\sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}};\end{aligned}$$

thus with C denoting the constant $\frac{2}{L}\|v_0\|_{L^1(0,L)} e^{\frac{\pi^2}{L^2}}\sum\limits_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}}$, we have

$$\sup_{x\in[0,L]}\Big|u(x,t) - \frac{b-a}{L}x - a\Big| \leqslant Ce^{-\frac{\pi^2}{L^2}t}. \tag{10.2.7}$$

Since $C < \infty$, we conclude that the function $u(\cdot,t)$ converges to the function $\dfrac{b-a}{L}x + a$ uniformly on $[0,L]$ as $t \to \infty$.

The Neumann problem

In this sub-section we consider the heat equation with Neumann boundary condition:

$$u_t - u_{xx} = 0 \qquad \text{in} \quad (0,L)\times(0,\infty), \tag{10.2.8a}$$

$$u = u_0 \qquad \text{on} \quad (0,L)\times\{t=0\}, \tag{10.2.8b}$$

$$u_x(0,t) = a, \quad u_x(L,t) = b \qquad \text{for all } t > 0, \tag{10.2.8c}$$

where a and b are given constants. Let $v(x,t) = u(x,t) - \dfrac{b-a}{2L}(x^2+2t) - ax$. Then v satisfies

$$\begin{aligned} v_t - v_{xx} &= 0 \qquad \text{in} \quad (0,L)\times(0,\infty)\,, && (10.2.9\text{a})\\ v &= v_0 \qquad \text{on} \quad (0,L)\times\{t=0\}\,, && (10.2.9\text{b})\\ v_x(0,t) = v_x(L,t) &= 0 \qquad \text{for all } t>0\,, && (10.2.9\text{c}) \end{aligned}$$

where $v_0 : [0,L] \to \mathbb{R}$ is given by $v_0(x) = u_0(x) - \dfrac{b-a}{2L}x^2 - ax$. As long as the solution v to (10.2.9) is found, the solution u to (10.2.8) can be constructed using $u(x,t) = v(x,t) + \dfrac{b-a}{2L}x^2 + ax$. Therefore, we focus on solving (10.2.9) (using the Fourier series method). We look for $\{d_k(t)\}_{k=0}^{\infty}$ such that

$$v(x,t) = \frac{d_0(t)}{2} + \sum_{k=1}^{\infty} d_k(t)\cos\frac{\pi k x}{L}$$

validates (10.2.9a,b).

Assume that the differentiation of the series can be obtained by term-by-term differentiation; that is,

$$\begin{aligned} \frac{\partial}{\partial t}\sum_{k=1}^{\infty} d_k(t)\cos\frac{\pi kx}{L} &= \sum_{k=1}^{\infty}\frac{\partial}{\partial t}\Big(d_k(t)\cos\frac{\pi kx}{L}\Big) = \sum_{k=1}^{\infty} d_k'(t)\cos\frac{\pi kx}{L}\,,\\ \frac{\partial}{\partial x}\sum_{k=1}^{\infty} d_k(t)\cos\frac{\pi kx}{L} &= \sum_{k=1}^{\infty}\frac{\partial}{\partial x}\Big(d_k(t)\cos\frac{\pi kx}{L}\Big) = -\sum_{k=1}^{\infty}\frac{k\pi}{L}d_k(t)\sin\frac{\pi kx}{L}\,, \end{aligned}$$

and

$$\frac{\partial^2}{\partial x^2}\sum_{k=1}^{\infty} d_k(t)\cos\frac{\pi kx}{L} = \sum_{k=1}^{\infty}\frac{\partial^2}{\partial x^2}\Big(d_k(t)\cos\frac{\pi kx}{L}\Big) = -\sum_{k=1}^{\infty}\frac{k^2\pi^2}{L^2}d_k(t)\cos\frac{\pi kx}{L}\,.$$

Then (10.2.9c) holds automatically, and (10.2.9a) implies that

$$\frac{d_0'(t)}{2} + \sum_{k=1}^{\infty}\Big[d_k'(t) + \frac{k^2\pi^2}{L^2}d_k(t)\Big]\cos\frac{\pi kx}{L} = 0\,.$$

Therefore, d_0 is a constant and d_k satisfies (10.2.5) as well. Moreover, expressing v_0 in terms of cosine series, (10.2.9b) implies that

$$d_k(0) = \widehat{v}_{0k} \equiv \frac{2}{L}\int_0^L v_0(x)\cos\frac{\pi kx}{L}\,dx \qquad \forall\, k\in\mathbb{N}\cup\{0\}\,.$$

Solving (10.2.5) with the initial condition above, we obtain that

$$d_k(t) = \widehat{v}_{0k}e^{-\frac{k^2\pi^2}{L^2}t} \qquad \forall\, k\in\mathbb{N};$$

thus the solution to (10.2.9) can be written as

$$v(x,t) = \frac{1}{L}\int_0^L v_0(x)\,dx + \sum_{k=1}^{\infty} \widehat{v}_{0k} e^{-\frac{k^2\pi^2}{L^2}t} \cos\frac{k\pi x}{L}\,.$$

Therefore, the solution to (10.2.3) can be written as

$$u(x,t) = \frac{1}{L}\int_0^L v_0(x)\,dx + \sum_{k=1}^{\infty} \widehat{v}_{0k} e^{-\frac{k^2\pi^2}{L^2}t} \cos\frac{k\pi x}{L} + \frac{b-a}{2L}(x^2+2t) + ax\,.$$

• the long time behavior: Suppose that the rod is insulated at the end-points; that is, the temperature u satisfies $u_x(0,t) = u_x(L,t) = 0$ for all $t > 0$. Then $v_0 = u_0$ and we expect that no matter what the temperature distribution is given initially, the temperature distribution approaches the average temperature; that is, we expect that $u(x,t) \to \dfrac{1}{L}\displaystyle\int_0^L u_0(x)\,dx$ as $t \to \infty$ for all $x \in [0,L]$. Similar to the derivation of (10.2.7),

$$\begin{aligned}\Big|u(x,t) - \frac{1}{L}\int_0^L u_0(x)\,dx\Big| &\leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)} \sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}t} \leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)} \sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}(t-1)} e^{-\frac{k^2\pi^2}{L^2}} \\ &\leqslant \frac{2}{L}\|v_0\|_{L^1(0,L)} e^{-\frac{\pi^2}{L^2}(t-1)} \sum_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}}\,;\end{aligned}$$

thus with C denoting the constant $\dfrac{2}{L}\|v_0\|_{L^1(0,L)} e^{\frac{\pi^2}{L^2}} \sum\limits_{k=1}^{\infty} e^{-\frac{k^2\pi^2}{L^2}}$, we have

$$\sup_{x\in[0,L]} \Big|u(x,t) - \frac{1}{L}\int_0^L u_0(x)\,dx\Big| \leqslant Ce^{-\frac{\pi^2}{L^2}t}. \tag{10.2.10}$$

Since $C < \infty$, we conclude that the function $u(\cdot,t)$ converges to the function $\dfrac{1}{L}\displaystyle\int_0^L u_0(x)\,dx$ uniformly on $[0,L]$ as $t \to \infty$.

10.2.2 Heat Conduction on $\mathbb{R}^n$

Consider the heat equation on $\mathbb{R}^n$

$$u_t - \Delta u = 0 \qquad \text{in} \quad \mathbb{R}^n \times (0,\infty), \tag{10.2.11a}$$

$$u = u_0 \qquad \text{on} \quad \mathbb{R}^n \times \{t=0\}, \tag{10.2.11b}$$

where Δ is the Laplace operator, called Laplacian, defined by

$$\Delta u = \sum_{k=1}^{n} \frac{\partial^2 u}{\partial x_k^2}\,(= \mathrm{div}\nabla u)\,.$$

For a function f of x (and probably also t), let $\mathscr{F}(f) = \widehat{f}$ denote the Fourier transform of f in x; that is,

$$\mathscr{F}(f)(\xi,t) = \widehat{f}(\xi,t) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} f(x,t)e^{-ix\cdot\xi}\,dx\,.$$

Then by assuming that $\widehat{u}(\cdot,t)$ exists for all $t>0$, Lemma 9.11 implies that

$$\begin{aligned}\mathscr{F}(\Delta u)(\xi,t) &= \mathscr{F}\Big(\sum_{k=1}^n \frac{\partial}{\partial x_k}\frac{\partial u}{\partial x_k}\Big)(\xi,t) = \sum_{k=1}^n \mathscr{F}\Big(\frac{\partial}{\partial x_k}\frac{\partial u}{\partial x_k}\Big)(\xi,t)\\ &= \sum_{k=1}^n i\xi_k\mathscr{F}\Big(\frac{\partial u}{\partial x_k}\Big)(\xi,t) = \sum_{k=1}^n (i\xi_k)^2\widehat{u}(\xi,t) = -|\xi|^2\widehat{u}(\xi,t)\,.\end{aligned}$$

Assume further that

$$\begin{aligned}\frac{\partial}{\partial t}\widehat{u}(\xi,t) &= \frac{\partial}{\partial t}\frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} u(x,t)e^{-ix\cdot\xi}\,dx = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n}\frac{\partial}{\partial t}u(x,t)e^{-ix\cdot\xi}\,dx\\ &= \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} u_t(x,t)e^{-ix\cdot\xi}\,dx = \widehat{u_t}(\xi,t)\,.\end{aligned}$$

Taking the Fourier transform of (10.2.11a), we find that

$$\frac{\partial}{\partial t}\widehat{u}(\xi,t) + |\xi|^2\widehat{u}(\xi,t) = \mathscr{F}(u_t - \Delta u)(\xi,t) = 0\,. \tag{10.2.12}$$

Since $\widehat{u}(\xi,0) = \mathscr{F}(u(\cdot,0))(\xi) = \widehat{u_0}(\xi)$, solving the ODE (10.2.12) with this initial condition we obtain that

$$\widehat{u}(\xi,t) = \widehat{u_0}(\xi)e^{-|\xi|^2 t} = \widehat{u_0}(\xi)\widehat{P_{2t}}(\xi)\,,$$

where $P_t(x) = \dfrac{1}{\sqrt{t}^n}e^{-\frac{|x|^2}{2t}}$. By Theorem 9.26, we conclude that

$$u(x,t) = (u_0 * P_{2t})(x) = \frac{1}{\sqrt{4\pi t}^n}\int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}}u_0(y)\,dy\,. \tag{10.2.13}$$

This induces the following

Definition 10.7. The function $\mathcal{H}(x,t) = \dfrac{1}{\sqrt{4\pi t}^n}e^{-\frac{|x|^2}{4t}}$ is called the ***heat kernel***.

Having introduced the heat kernel, the solution to (10.2.11), given by (10.2.13) can be expressed by

$$u(x,t) = \big(\mathcal{H}(\cdot,t) * u_0\big)(x)\,.$$

• Non-uniqueness of solutions: The Fourier transform method only picks up solutions whose Fourier transform is defined, and it is possible that there are other solutions to (10.2.11). Consider the function

$$u(x,t)=\sum_{k=0}^{\infty}\frac{g^{(k)}(t)}{(2k)!}x^{2k}, \tag{10.2.14}$$

where g is given by

$$g(t)=\begin{cases}\exp(-t^{-2}) & \text{if } t>0,\\ 0 & \text{if } t=0.\end{cases}$$

Then there exists $\theta>0$ such that

$$\big|g^{(k)}(t)\big|\leqslant\frac{k!}{(\theta t)^k}\exp\big(-\frac{1}{2}t^{-2}\big)\qquad\forall\, t>0. \tag{10.2.15}$$

In fact, using the Cauchy integral formula,

$$g^{(k)}(t)=\frac{k!}{2\pi i}\oint_{|z-t|=\theta t}\frac{g(z)}{(z-t)^{k+1}}\,dz$$

where $\theta\in(0,1)$ is chosen so small such that $\big|g(z)\big|\leqslant\exp\big(-\frac{1}{2}t^{-2}\big)$ on $|z-t|=\theta t$. The choice of such a θ is possible since by writing $z=x+iy$,

$$\big|g(x+iy)\big|=\Big|\exp\big(\frac{y^2-x^2+2ixy}{(x^2+y^2)^2}\big)\Big|=\exp\big(\frac{y^2-x^2}{(x^2+y^2)^2}\big)\leqslant\exp\big(\frac{\theta^2-(1-\theta)^2}{((1+\theta)^2+\theta^2)^2}t^{-2}\big)\text{ if }\theta<\frac{1}{2}.$$

By the fact that $\dfrac{k!}{(2k)!}\leqslant\dfrac{1}{k!}$, we find that

$$\sum_{k=0}^{\infty}\Big|\frac{g^{(k)}(t)}{(2k)!}x^{2k}\Big|\leqslant\sum_{k=0}^{\infty}\frac{x^{2k}}{k!(\theta t)^k}\exp\big(-\frac{1}{2}t^{-2}\big)=\exp\Big[\frac{1}{t}\big(\frac{x^2}{\theta}-\frac{1}{2}t^{-2}\big)\Big]\quad\forall\, t>0, x\in\mathbb{R}. \tag{10.2.16}$$

Therefore, by comparison the series in (10.2.14) converges for all $t>0$, and trivially also converges for $t=0$. Moreover, (10.2.16) shows that for each $t\in\mathbb{R}$, (10.2.14) is a convergent power series; thus

$$u_{xx}(x,t)=\frac{\partial^2}{\partial x^2}u(x,t)=\sum_{k=0}^{\infty}\frac{g^{(k)}(t)}{(2k)!}\frac{\partial^2 x^{2k}}{\partial x^2}=\sum_{k=1}^{\infty}\frac{g^{(k)}(t)}{(2k-2)!}x^{2k-2}\qquad\forall\, t>0, x\in\mathbb{R}.$$

On the other hand, by the fact that $\dfrac{(k+1)!}{(2k)!}\leqslant\dfrac{1}{(k-1)!}$ if $k\geqslant 1$, using (10.2.15) we find that for all $t>0$ and $x\in\mathbb{R}$,

$$\begin{aligned}\sum_{k=0}^{\infty}\Big|\frac{g^{(k+1)}(t)}{(2k)!}x^{2k}\Big|&\leqslant\big|g'(t)\big|+\sum_{k=1}^{\infty}\frac{x^{2k}}{(k-1)!(\theta t)^{k+1}}\exp\big(-\frac{1}{2}t^{-2}\big)\\&=\big|g'(t)\big|+\frac{x^2}{(\theta t)^2}\exp\Big[\frac{1}{t}\big(\frac{x^2}{\theta}-\frac{1}{2}t^{-2}\big)\Big].\end{aligned}$$

Therefore, the series $\sum\limits_{k=0}^{\infty} \dfrac{g^{(k+1)}(t)}{(2k)!} x^{2k}$ converges uniformly on any bounded set of $\mathbb{R}$; thus

$$u_t(x,t) = \sum_{k=0}^{\infty} \frac{g^{(k+1)}(t)}{(2k)!} x^{2k} = \sum_{k=1}^{\infty} \frac{g^{(k)}(t)}{(2k-2)!} x^{2k-2} = u_{xx}(x,t) \qquad \forall\, t > 0, x \in \mathbb{R}.$$

This implies that u satisfies the heat equation

$$\begin{aligned} u_t - u_{xx} &= 0 \qquad \text{in} \quad \mathbb{R} \times (0,\infty)\,, \\ u &= 0 \qquad \text{on} \quad \mathbb{R} \times \{t = 0\}\,. \end{aligned}$$

Note that using the Fourier transform method to solve the PDE above we obtain trivial solution. The reason for not seeing the solution given by (10.2.14) using the Fourier transform method is that the Fourier transform of the function u given by (10.2.14) does not exists.

10.3 Exercises

§10.1 Application on Signal Processing

In this part of exericse, the Fourier transform is always defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x\cdot\xi} dx \quad \text{and} \quad \check{f}(x) = \int_{\mathbb{R}^n} f(\xi) e^{2\pi i x\cdot\xi} d\xi \qquad \forall\, f \in L^1(\mathbb{R}^n)\,. \tag{10.1.1}$$

Problem 10.1. Let $f \in L^2(\mathbb{R})$ be given by $f(t) = \sin(at)\sin(bt)/t$, with a and b positive. Show that f is band-limited in the sense that $\widehat{f}(\xi) = 0$ for $|\xi| > a + b$. You may use that standard properties of the Fourier transform remain valid for L^2-functions, even though we have not gone into this is detail.

Problem 10.2. If δ and Ω are two numbers with $0 < \pi/\Omega < \delta$, find a function $f \in L^2(\mathbb{R})$, such that $\widehat{f}(\xi) = 0$ for $|\xi| \geqslant \Omega$ and $f(n\delta) = 0$ for $n \in \mathbb{Z}$, but $f \neq 0$ as an element of $L^2(\mathbb{R})$. This shows that the Shannon–Nyquist sampling distance $\Delta t = \pi/\xi_{\max}$ is the largest possible.
Hint: Use Problem 10.1.

Problem 10.3. Complete the following.

1. Show the Poisson summation formula
$$\sum_{k=-\infty}^{\infty} \widehat{f}(k) = \sum_{n=-\infty}^{\infty} f(n) \qquad \forall\, f \in \mathscr{S}(\mathbb{R})\,.$$

2. Suppose that $f \in \mathscr{S}(\mathbb{R})$. Show that

$$\sum_{k=-\infty}^{\infty} \widehat{f}(\xi - \frac{k}{T}) = T \sum_{n=-\infty}^{\infty} f(nT) e^{-i2\pi nT\xi} \qquad \forall\, \xi \in \mathbb{R}\,.$$

In particular, if $f \in \mathscr{S}(\mathbb{R})$ and $\operatorname{supp}(\widehat{f}) \subseteq \big[-\frac{1}{T}, \frac{1}{T}\big]$,

$$\widehat{f}(\xi) = T \sum_{n=-\infty}^{\infty} f(nT) e^{-i2\pi nT\xi} \qquad \forall\, \xi \in \big[-\frac{1}{T}, \frac{1}{T}\big]\,.$$

This implies that if $\widehat{f}$ has compact support in $\big[-\frac{1}{T}, \frac{1}{T}\big]$, f can be reconstructed based on partial knowledge of f, namely $\{f(nT)\}_{n=-\infty}^{\infty}$.

3. Show that the Poisson summation formula

$$\sum_{k=-\infty}^{\infty} \widehat{\phi}(k) = \sum_{n=-\infty}^{\infty} \phi(n)$$

also holds for function ϕ satisfying

$$\big|\phi(x)\big| + \big|\widehat{\phi}(x)\big| \leqslant C(1+|x|)^{-1-\delta} \qquad \forall\, x \in \mathbb{R}\,.$$

Hint of 1: For $f \in \mathscr{S}(\mathbb{R})$, define $F : \mathbb{R} \to \mathbb{R}$ by $F(x) = \sum\limits_{n=-\infty}^{\infty} f(n+x)$. Show that $F \in \mathscr{C}^1(\mathbb{R};\mathbb{R})$ and F is periodic with period 1. Express F in terms of Fourier series and find that value $F(0)$.

Hint of 2: For each $\xi \in \mathbb{R}$, let $f(x) = g(Tx)e^{-i2\pi Tx\xi}$ and use 1.

Hint of 3: Make use of the conclusion in Problem 8.11.

Problem 10.4. Let $\text{Ш} : \mathscr{S}(\mathbb{R}) \to \mathbb{C}$ be the Shah function, also called the Comb function, denoted by

$$\langle \text{Ш}, \phi \rangle = \sum_{n=-\infty}^{\infty} \phi(n) \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,.$$

1. Show that Ш is a tempered distribution.

2. Show that $\widehat{\text{Ш}} = \text{Ш}$ in $\mathscr{S}(\mathbb{R})'$; that is,

$$\langle \widehat{\text{Ш}}, \phi \rangle = \langle \text{Ш}, \phi \rangle \qquad \forall\, \phi \in \mathscr{S}(\mathbb{R})\,.$$

3. For $p > 0$, define $\text{Ш}_p = \dfrac{1}{p} d_p \text{Ш}$. Show that

$$\langle \text{Ш}_p, \phi \rangle = \sum_{n=-\infty}^{\infty} \phi(pn) \qquad \text{and} \qquad \widehat{\text{Ш}_p} = \frac{1}{p} \text{Ш}_{\frac{1}{p}} \,.$$

4. Show that if $\phi : \mathbb{R} \to \mathbb{C}$ is integrable (that is, $\phi \in L^1(\mathbb{R})$) and ϕ satisfies that $\operatorname{supp}(\phi) \subseteq (0, p)$ for some $p > 0$, then $\text{Ш}_p * \phi$ is a tempered distribution and is the periodic extension of ϕ with period p. In other words, for such a ϕ one has

$$(\text{Ш} * \phi)(x) = \phi(x) \quad \forall\, x \in (0, p) \qquad \text{and} \qquad (\text{Ш} * \phi)(x + np) = (\text{Ш} * \phi)(x) \quad \forall\, x \in \mathbb{R} \,.$$

Appendix A

The Inverse and Implicit Function Theorem

反函數定理是用來探討一個函數的反函數是否存在的問題。只要一個函數不是一對一的，一般來說都不能定義其反函數，例如三角函數中，正弦、餘弦及正切函數都是周期函數，所以全域的反函數不存在。但是我們也知道有所謂的反三角函數 $\sin^{-1}$（或 arcsin），$\cos^{-1}$（或 arctan）及 $\tan^{-1}$（或 arctan），這是因為我們限制了原三角函數的定義域使其在新的定義域上是一對一的（因此反函數存在）。因此，要討論一個定在某一個（大範圍的）定義域的函數的反函數，常常我們最多只能說反函數只在某一小塊區域上存在。

A.1 The Inverse Function Theorem for Functions of One Real Variable

如何知道一個函數在一小塊區域上的反函數存在，我們首先該問的是在定義域是一維（或是指單變數函數）的情況下發生什麼事？在開始敘述並證明單變數函數的反函數定理前，我們先複習一些微積分中的重要定理，尤其是均值定理。

Theorem A.1 (Rolle)**.** *Suppose that a function $f : [a, b] \to \mathbb{R}$ is continuous, and is differentiable on (a, b). If $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. By the Extreme Value Theorem (Corollary 4.26), there exists x_0 and x_1 in $[a, b]$ such that

$$f(x_0) = \min f([a, b]) \quad \text{and} \quad f(x_1) = \max f([a, b]) .$$

Case 1. $f(x_0) = f(x_1)$, then f is constant on $[a, b]$; thus $f'(x) = 0$ for all $x \in (a, b)$.

Case 2. One of $f(x_0)$ and $f(x_1)$ is different from $f(a)$. W.L.O.G. we may assume that $f(x_0) \neq f(a)$. Then $x_0 \in (a,b)$, and f attains its global minimum at x_0. Therefore, $f'(x_0) = 0$. □

Theorem A.2 (Cauchy's Mean Value Theorem)**.** *Suppose that functions $f, g : [a,b] \to \mathbb{R}$ are continuous, and $f, g : (a,b) \to \mathbb{R}$ are differentiable. If $g(a) \neq g(b)$ and $g'(x) \neq 0$ for all $x \in (a,b)$, then there exists $c \in (a,b)$ such that*

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Proof. Consider the function

$$h(x) \equiv \big(f(x) - f(a)\big)\big(g(b) - g(a)\big) - \big(f(b) - f(a)\big)\big(g(x) - g(a)\big).$$

Then $h : [a,b] \to \mathbb{R}$ is continuous, and is differentiable on (a,b). Moreover, $h(b) = h(a) = 0$. By Rolle's theorem, there exists $c \in (a,b)$ such that

$$h'(c) = f'(c)\big(g(b) - g(a)\big) - \big(f(b) - f(a)\big)g'(c) = 0.$$ □

Corollary A.3 (Mean Value Theorem)**.** *Suppose that a function $f : [a,b] \to \mathbb{R}$ is continuous, and $f : (a,b) \to \mathbb{R}$ is differentiable. Then there exists $c \in (a,b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof. Apply the Cauchy Mean Value Theorem for the case that $g(x) = x$. □

以下定理是單變數函數的反函數定理。

Theorem A.4 (Inverse Function Theorem)**.** *Let $f : (a,b) \to \mathbb{R}$ be differentiable, and f' be sign-definite; that is, $f'(x) > 0$ for all $x \in (a,b)$ or $f'(x) < 0$ for all $x \in (a,b)$. Then $f : (a,b) \to f((a,b))$ is a bijection, and f^{-1}, the inverse function of f, is differentiable on $f((a,b))$, and*

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)} \qquad \forall\, x \in (a,b). \tag{A.1.1}$$

Proof. W.L.O.G. we assume that $f'(x) > 0$ for all $x \in (a,b)$. Then f is strictly increasing; thus f^{-1} exists.

Claim: $f^{-1} : f((a,b)) \to (a,b)$ is continuous.

Proof of claim: Let $y_0 = f(x_0) \in f((a,b))$, and $\varepsilon > 0$ be given. Then $f((x_0-\varepsilon, x_0+\varepsilon)) = \big(f(x_0-\varepsilon), f(x_0+\varepsilon)\big)$ since f is continuous on (a,b) and $(x_0-\varepsilon, x_0+\varepsilon)$ is connected. Let $\delta = \min\{f(x_0)-f(x_0-\varepsilon), f(x_0+\varepsilon)-f(x_0)\big\}$. Then $\delta > 0$, and

$$(y_0-\delta, y_0+\delta) = \big(f(x_0)-\delta, f(x_0)+\delta\big) \subseteq f((x_0-\varepsilon, x_0+\varepsilon));$$

thus by the injectivity of f,

$$\begin{aligned} &f^{-1}((y_0-\delta, y_0+\delta)) \\ &\qquad \subseteq f^{-1}(f((x_0-\varepsilon, x_0+\varepsilon))) = (x_0-\varepsilon, x_0+\varepsilon) = (f^{-1}(y_0)-\varepsilon, f^{-1}(y_0)+\varepsilon). \end{aligned}$$

The inclusion above implies that f^{-1} is continuous at y_0.

Writing $y = f(x)$ and $x = f^{-1}(y)$. Then if $y_0 = f(x_0) \in f((a,b))$,

$$\frac{f^{-1}(y)-f^{-1}(y_0)}{y-y_0} = \frac{x-x_0}{f(x)-f(x_0)}.$$

Since f^{-1} is continuous on $f((a,b))$, $x \to x_0$ as $y \to y_0$; thus

$$\lim_{y\to y_0} \frac{f^{-1}(y)-f^{-1}(y_0)}{y-y_0} = \lim_{x\to x_0} \frac{x-x_0}{f(x)-f(x_0)} = \frac{1}{f'(x_0)}$$

which implies that f^{-1} is differentiable at y_0. □

A.2 Some Additional Tools in Analysis

A.2.1 The contraction mapping principle

Definition A.5. Let (M,d) be a metric space, and $\Phi : M \to M$ be a mapping. Φ is said to be a ***contraction mapping*** if there exists a constant $k \in [0,1)$ such that

$$d\big(\Phi(x), \Phi(y)\big) \leqslant k d(x,y) \qquad \forall\, x,y \in M.$$

Remark A.6. A contraction mapping must be (uniformly) continuous. To see this, let $\varepsilon > 0$ be given. Take $\delta = \dfrac{\varepsilon}{k}$, where k is set as in the definition of contraction. Now if $d(x,y) < \delta$, then

$$d\big(\Phi(x), \Phi(y)\big) \leqslant k d(x,y) < k \cdot \frac{\varepsilon}{k} = \varepsilon.$$

Example A.7. For what $r < 1$ do we have $f : [0, r] \to [0, r]$ where $f(x) = x^2$ a contraction?
Answer: By the mean value theorem, $f(x) - f(y) = f'(c)(x - y)$, c between x, y; thus

$$\big|f(x) - f(y)\big| = \big|f'(c)\big||x - y| = 2c|x - y| \leqslant 2r|x - y|\,.$$

Hence for all $r < \dfrac{1}{2}$, the map $f : [0, r] \to [0, r]$ is a contraction where $f(x) = x^2$.

On the other hand, we show that f cannot be a contraction if $r = \dfrac{1}{2}$. Suppose the contrary that there exists $k \in [0, 1)$ such that for all $x, y \in \big[0, \dfrac{1}{2}\big]$, $\big|x^2 - y^2\big| \leqslant k\big|x - y\big|$. Then

$$\sup_{x\neq y, x,y\in[0,\frac{1}{2}]} \frac{\big|x^2 - y^2\big|}{\big|x - y\big|} \leqslant k < 1\,.$$

But we can take $x = \dfrac{1}{2}$, $y_n = \dfrac{1}{2} - \dfrac{1}{n}$, $n = 1, 2, \cdots, x, y_n \in \big[0, \dfrac{1}{2}\big]$. So

$$\lim_{n\to\infty} \frac{\big|x^2 - {y_n}^2\big|}{\big|x - y_n\big|} = \lim_{n\to\infty} \big|x + y_n\big| = \lim_{n\to\infty} \Big(\frac{1}{2} + \frac{1}{2} - \frac{1}{n}\Big) = 1\,.$$

This means $\displaystyle\sup_{x\neq y, x,y\in[0,\frac{1}{2}]} \frac{\big|x^2 - y^2\big|}{\big|x - y\big|} < 1$ is not possible.

Definition A.8. Let (M, d) be a metric space, and $\Phi : M \to M$ be a mapping. A point $x_0 \in M$ is called a ***fixed-point*** for Φ if $\Phi(x_0) = x_0$.

Example A.9. Let $\Phi : \mathbb{R} \to \mathbb{R}$ be given by $\Phi(x) = \dfrac{x^2 + 2}{3}$. Then 1 is a fixed-point, and 2 is also a fixed-point.

Theorem A.10 (Contraction Mapping Principle)**.** *Let (M, d) be a complete metric space, and $\Phi : M \to M$ be a contraction mapping. Then Φ has a unique fixed-point.*

Proof. Let $x_0 \in M$, and define $x_{n+1} = \Phi(x_n)$ for all $n \in \mathbb{N} \cup \{0\}$. Then

$$d(x_{n+1}, x_n) = d\big(\Phi(x_n), \Phi(x_{n-1})\big) \leqslant kd(x_n, x_{n-1}) \leqslant k^n d(x_1, x_0)\,;$$

thus if $n > m$,

$$\begin{aligned} d(x_n, x_m) &\leqslant d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \cdots + d(x_{n-1}, x_n) \\ &\leqslant (k^m + k^{m+1} + \cdots + k^{n-1})d(x_1, x_0) \\ &\leqslant k^m(1 + k + k^2 + \cdots)d(x_1, x_0) = \frac{k^m}{1-k}d(x_1, x_0)\,. \end{aligned} \tag{A.2.1}$$

Since $k \in [0,1)$, $\lim\limits_{m\to\infty} \dfrac{k^m}{1-k} d(x_1,x_0) = 0$; thus

$$\forall\, \varepsilon > 0, \exists\, N > 0 \ni d(x_n,x_m) < \varepsilon \quad \forall\, n,m \geqslant N\,.$$

In other words, $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence. Since (M,d) is complete, $x_n \to x$ as $n\to\infty$ for some $x \in M$. Finally, since $\Phi(x_n) = x_{n+1}$ for all $n \in \mathbb{N}$, by the continuity of Φ we obtain that

$$\Phi(x) = \lim_{n\to\infty} \Phi(x_n) = \lim_{n\to\infty} x_{n+1} = x$$

which guarantees the existence of a fixed-point.

Suppose that for some $x, y \in M$, $\Phi(x) = x$ and $\Phi(y) = y$. Then

$$d(x,y) = d\big(\Phi(x),\Phi(y)\big) \leqslant k d(x,y)$$

which implies that $d(x,y) = 0$ or $x = y$. Therefore, the fixed-point of Φ is unique. □

Remark A.11. The proof of the contraction mapping principle also provides an iterative way, $x_{k+1} = \Phi(x_k)$, of finding the fixed-point of a contraction mapping Φ. Using (A.2.1), the convergence rate of $\{x_m\}_{m=1}^\infty$ to the fixed-point x is measured by

$$d(x_m,x) = \lim_{n\to\infty} d(x_m,x_n) \leqslant \frac{k^m}{1-k} d(x_1,x_0)\,.$$

Therefore, the smaller the contraction constant k, the faster the convergence.

Remark A.12. Theorem A.10 sometimes is also called the ***Banach fixed-point theorem***.

Example A.13. The condition $k < 1$ in Theorem A.10 is necessary. For example, let $M = \mathbb{R}$, $d(x,y) = |x-y|$, and $\Phi : \mathbb{R} \to \mathbb{R}$ be given by $\Phi(x) = x+1$. Then $\big|\Phi(x) - \Phi(y)\big| = \big|x-y\big|$. Suppose $x_\star$ is a fixed-point of Φ. Then $x_\star = \Phi(x_\star) = x_\star + 1$ which leads to a contradiction that $0 = 1$.

Example A.14. Let $\Phi : [1,\infty) \to [1,\infty)$ be given by $\Phi(x) = x + \dfrac{1}{x}$. Then if $x \neq y$,

$$\big|\Phi(x) - \Phi(y)\big| = \Big|x - y + \frac{1}{x} - \frac{1}{y}\Big| = \Big|(x-y)\Big(1 - \frac{1}{xy}\Big)\Big| < |x-y|\,.$$

However, there is no fixed-point of Φ.

Example A.15 (The secant method)**.** Suppose that f is continuously differentiable, $f'(x) > 0$ for all $x \in [a,b]$ and $f(a)f(b) < 0$. By the intermediate value theorem there must be a (unique) zero of f. How do we find this zero?

Assume that $\sup\limits_{x\in[a,b]} f'(x) < \infty$. Let

$$M = \max\Big\{\sup_{x\in[a,b]} f'(x),\ -\frac{f(a)}{b-a}, \frac{f(b)}{b-a}\Big\} + 1$$

be a positive constant, and consider $\Phi(x) = x - \dfrac{f(x)}{M}$. Then by the mean value theorem,

$$\big|\Phi(x) - \Phi(y)\big| = \Big|(x-y)(1 - \frac{f'(\xi)}{M})\Big| \leqslant \big(1 - \frac{\min_{\xi\in[a,b]} f'(\xi)}{M}\big)|x-y| \leqslant k|x-y|\,,$$

where $k \in [0,1)$ is a fixed constant. Moreover, $\Phi'(x) = 1 - \dfrac{f'(x)}{M} > 0$; thus Φ is strictly increasing. Since the choice of M implies that $a < \Phi(a) < \Phi(b) < b$; thus $\Phi : [a,b] \to [a,b]$. Therefore, the contraction mapping principle implies that one can find the fixed-point of Φ (which is the zero of f) using the iterative scheme $x_{k+1} = \Phi(x_k)$ (by picking any arbitrary initial guess $x_0 \in [a,b]$).

A.2.2 The mean value theorem for functions of several variables

Theorem A.16. *Let $U \subseteq \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}^m$ with $f = (f_1, \cdots, f_m)$. Suppose that f is differentiable on U and the line segment joining x and y lies in U. Then there exist points $c_1, \cdots, c_m$ on that segment such that*

$$f_i(y) - f_i(x) = (Df_i)(c_i)(y-x) \qquad \forall\, i = 1, \cdots, m.$$

Moreover, if U is convex and $\sup\limits_{x\in U} \|(Df)(x)\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} \leqslant M$, then

$$\|f(x) - f(y)\|_{\mathbb{R}^m} \leqslant M\|x-y\|_{\mathbb{R}^n} \qquad \forall\, x, y \in U\,.$$

Proof. Let $\gamma : [0,1] \to \mathbb{R}^n$ be given by $\gamma(t) = (1-t)x + ty$. Then by Theorem 5.47, for each $i = 1, \cdots, m$, $(f_i \circ \gamma) : [0,1] \to \mathbb{R}$ is differentiable on $(0,1)$; thus the mean value theorem (Corollary A.3) implies that there exists $t_i \in (0,1)$ such that

$$f_i(y) - f_i(x) = (f_i\circ\gamma)(1) - (f_i\circ\gamma)(0) = (f_i\circ\gamma)'(t_i) = (Df_i)(c_i)\big(\gamma'(t_i)\big)\,,$$

where $c_i = \gamma(t_i)$. On the other hand, $\gamma'(t_i) = y - x$.

Let $g(t) = (f \circ \gamma)(t)$. Then the chain rule implies that $g'(t) = (Df)(\gamma(t))(y-x)$; thus

$$\|g'(t)\|_{\mathbb{R}^m} \leqslant \|(Df)(\gamma(t))\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} \|y-x\|_{\mathbb{R}^m} \leqslant M\|x-y\|_{\mathbb{R}^n}.$$

Define $h(t) = \big(g(1)-g(0)\big) \cdot g(t)$. Then $h : [0,1] \to \mathbb{R}$ is differentiable; thus by the mean value theorem we find that there exists $\xi \in (0,1)$ such that

$$h(1) - h(0) = h'(\xi) = \big(g(1)-g(0)\big) \cdot g'(\xi);$$

thus by the fact that $g(0) = f(x)$ and $g(1) = f(y)$,

$$\begin{aligned}\|f(x)-f(y)\|_{\mathbb{R}^m}^2 = h(1)-h(0) &\leqslant \|g(1)-g(0)\|_{\mathbb{R}^m}\|g'(\xi)\|_{\mathbb{R}^m} \\ &\leqslant M\|f(x)-f(y)\|_{\mathbb{R}^m}\|x-y\|_{\mathbb{R}^n}\end{aligned}$$

which concludes the theorem. □

A.3 The Inverse Function Theorem（反函數定理）

接下來我們探討多變數函數的反函數定理。由一維的反函數定理 (Theorem A.4) 我們知道首先應該要保留的條件是類似於微分不為零的這個條件。但是在多變數函數之下，微分不為零的條件該怎麼呈現，這是第一個問題。而當我們觀察 (A.1.1)，應該可以猜出在多變數版本裡面所該對應到的條件，即是 $(Df)(x)$ 這個 bounded linear map 的可逆性。

另外，假設 $f \in \mathscr{C}^1$，那麼由 Theorem 5.11 我們知道在一個點 x_0 如果 $(Df)(x_0)$ 可逆的話，那麼在一個鄰域裡 $(Df)(x)$ 都可逆。所以下面這個反函數定理的條件中只有 (Df) 在一個點可逆這個條件，因為我們目前想先知道小區域的反函數存不存在。

Theorem A.17 (Inverse Function Theorem)**.** *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open, $x_0 \in \mathcal{D}$, $f : \mathcal{D} \to \mathbb{R}^n$ be of class $\mathscr{C}^1$, and $(Df)(x_0)$ be invertible. Then there exist an open neighborhood U of x_0 and an open neighborhood V of $f(x_0)$ such that*

1. *$f : U \to V$ is one-to-one and onto;*
2. *The inverse function $f^{-1} : V \to U$ is of class $\mathscr{C}^1$;*
3. *If $x = f^{-1}(y)$, then $(Df^{-1})(y) = \big((Df)(x)\big)^{-1}$;*
4. *If f is of class $\mathscr{C}^r$ for some $r > 1$, so is f^{-1}.*

Proof. Assume that $A = (Df)(x_0)$. Then $\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} \neq 0$. Choose $\lambda > 0$ such that $2\lambda\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} = 1$. Since $f \in \mathscr{C}^1$, there exists $\delta > 0$ such that

$$\big\|(Df)(x) - A\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} = \big\|(Df)(x) - (Df)(x_0)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} < \lambda \quad \text{whenever } x \in B(x_0,\delta) \cap \mathcal{D}.$$

By choosing δ even smaller if necessary, we can assume that $B(x_0,\delta) \subseteq \mathcal{D}$. Let $U = B(x_0,\delta)$.
Claim: $f : U \to \mathbb{R}^n$ is one-to-one (hence $f : U \to f(U)$ is one-to-one and onto).
Proof of claim: For each $y \in \mathbb{R}^n$, define $\varphi_y(x) = x + A^{-1}\big(y - f(x)\big)$ (and we note that every fixed-point of φ_y corresponds to a solution to $f(x) = y$). Then

$$(D\varphi_y)(x) = \mathrm{Id} - A^{-1}(Df)(x) = A^{-1}\big(A - (Df)(x)\big),$$

where Id is the identity map on $\mathbb{R}^n$. Therefore,

$$\big\|(D\varphi_y)(x)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} \leqslant \|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\big\|A - (Df)(x)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} < \frac{1}{2} \quad \forall\, x \in B(x_0,\delta).$$

By the mean value theorem (Theorem A.16),

$$\big\|\varphi_y(x_1) - \varphi_y(x_2)\big\|_{\mathbb{R}^n} \leqslant \frac{1}{2}\|x_1 - x_2\|_{\mathbb{R}^n} \qquad \forall\, x_1, x_2 \in B(x_0,\delta),\ x_1 \neq x_2; \tag{A.3.1}$$

thus at most one x satisfies $\varphi_y(x) = x$; that is, φ_y has at most one fixed-point. As a consequence, $f : B(x_0,\delta) \to \mathbb{R}^n$ is one-to-one.
Claim: The set $V = f(U)$ is open.
Proof of claim: Let $b \in V$. Then there is $a \in U$ with $f(a) = b$. Choose $r > 0$ such that $\overline{B(a,r)} \subseteq U$. We observe that if $y \in B(b,\lambda r)$, then

$$\|\varphi_y(a) - a\|_{\mathbb{R}^n} \leqslant \|A^{-1}\big(y - f(a)\big)\|_{\mathbb{R}^n} \leqslant \|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\|y - b\|_{\mathbb{R}^n} < \lambda\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} r = \frac{r}{2};$$

thus if $y \in B(b,\lambda r)$ and $x \in B(a,r)$,

$$\|\varphi_y(x) - a\|_{\mathbb{R}^n} \leqslant \big\|\varphi_y(x) - \varphi_y(a)\big\|_{\mathbb{R}^n} + \|\varphi_y(a) - a\|_{\mathbb{R}^n} < \frac{1}{2}\|x - a\|_{\mathbb{R}^n} + \frac{r}{2} < r.$$

Therefore, if $y \in B(b,\lambda r)$, then $\varphi_y : B(a,r) \to B(a,r)$. By the continuity of φ_y,

$$\varphi_y : \overline{B(a,r)} \to \overline{B(a,r)}.$$

On the other hand, (A.3.1) implies that φ_y is a contraction mapping if $y \in B(b,\lambda r)$; thus by the contraction mapping principle A.10 φ_y has a unique fixed-point $x \in B(a,r)$. As a result, every $y \in B(b, \lambda r)$ corresponds to a unique $x \in B(a, r)$ such that $\varphi_y(x) = x$ or equivalently, $f(x) = y$. Therefore,

$$B(b, \lambda r) \subseteq f\big(B(a, r)\big) \subseteq f(U) = V\,.$$

Next we show that $f^{-1} : V \to U$ is differentiable. We note that if $x \in B(x_0, \delta)$,

$$\|(Df)(x_0) - (Df)(x)\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} < \lambda\|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)} = \frac{1}{2};$$

thus Theorem 5.11 implies that $(Df)(x)$ is invertible if $x \in B(x_0, \delta)$.

Let $b \in V$ and $k \in \mathbb{R}^n$ such that $b + k \in V$. Then there exists a unique $a \in U$ and $h = h(k) \in \mathbb{R}^n$ such that $a + h \in U$, $b = f(a)$ and $b + k = f(a + h)$. By the mean value theorem and (A.3.1),

$$\big\|\varphi_y(a + h) - \varphi_y(a)\big\|_{\mathbb{R}^n} < \frac{1}{2}\|h\|_{\mathbb{R}^n};$$

thus the fact that $f(a + h) - f(a) = k$ implies that

$$\|h - A^{-1}k\|_{\mathbb{R}^n} < \frac{1}{2}\|h\|_{\mathbb{R}^n}$$

which further shows that

$$\frac{1}{2}\|h\|_{\mathbb{R}^n} \leqslant \|A^{-1}k\|_{\mathbb{R}^n} \leqslant \|A^{-1}\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\|k\|_{\mathbb{R}^n} \leqslant \frac{1}{2\lambda}\|k\|_{\mathbb{R}^n}\,. \tag{A.3.2}$$

As a consequence, if k is such that $b + k \in V$,

$$\begin{aligned}
\frac{\big\|f^{-1}(b + k) - f^{-1}(b) - \big((Df)(a)\big)^{-1}k\big\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} &= \frac{\big\|a + h - a - \big((Df)(a)\big)^{-1}k\big\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} \\
&\leqslant \big\|\big((Df)(a)\big)^{-1}\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\frac{\big\|k - (Df)(a)(h)\big\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} \\
&\leqslant \big\|\big((Df)(a)\big)^{-1}\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}\frac{\big\|f(a + h) - f(a) - (Df)(a)(h)\big\|_{\mathbb{R}^n}}{\|h\|_{\mathbb{R}^n}}\frac{\|h\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} \\
&\leqslant \frac{\big\|\big((Df)(a)\big)^{-1}\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)}}{\lambda}\frac{\big\|f(a + h) - f(a) - (Df)(a)(h)\big\|_{\mathbb{R}^n}}{\|h\|_{\mathbb{R}^n}}\,.
\end{aligned}$$

Using (A.3.2), $h \to 0$ as $k \to 0$; thus passing $k \to 0$ on the left-hand side of the inequality above, by the differentiability of f we conclude that

$$\lim_{k \to 0}\frac{\big\|f^{-1}(b + k) - f^{-1}(b) - \big((Df)(a)\big)^{-1}k\big\|_{\mathbb{R}^n}}{\|k\|_{\mathbb{R}^n}} = 0\,.$$

This proves 3.

To see 4, we note that the map $g : \mathrm{GL}(n) \to \mathrm{GL}(n)$ given by $g(L) = L^{-1}$ is infinitely many time differentiable; thus using the identity

$$(Df^{-1})(y) = \big((Df)(x)\big)^{-1} = \big(g \circ (Df) \circ f^{-1}\big)(y),$$

by the chain rule we find that if $f \in \mathscr{C}^r$, then $Df^{-1} \in \mathscr{C}^{r-1}$ which is the same as saying that $f^{-1} \in \mathscr{C}^r$. □

Remark A.18. Since $f^{-1} : V \to U$ is continuous, for any open subset W of U $f(W) = (f^{-1})^{-1}(W)$ is open relative to V, or $f(W) = \mathcal{O} \cap V$ for some open set $\mathcal{O} \subseteq \mathbb{R}^n$. In other words, if U is an open neighborhood of x_0 given by the inverse function theorem, then $f(W)$ is also open for all open subsets W of U. We call this property as f is a ***local open mapping*** at x_0.

Remark A.19. Since $(Df)(x_0) \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$, the condition that $(Df)(x_0)$ is invertible can be replaced by that the determinant of the Jacobian matrix of f at x_0 is not zero; that is,

$$\det\big(\big[(Df)(x_0)\big]\big) \neq 0\,.$$

The determinant of the Jacobian matrix of f at x_0 is called the ***Jacobian*** of f at x_0. The Jacobian of f at x sometimes is denoted by $\mathbf{J}_f(x)$ or $\dfrac{\partial(f_1, \cdots, f_n)}{\partial(x_1, \cdots, x_n)}$.

Example A.20. Let $f : \mathbb{R} \to \mathbb{R}$ be given by

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

Let $0 \in (a, b)$ for some (small) open interval (a, b). Since $f'(x) = 1 - 2\cos\dfrac{1}{x} + 4x \sin\dfrac{1}{x}$ for $x \neq 0$, f has infinitely many critical points in (a, b), and (for whatever reasons) these critical points are local maximum points or local minimum points of f which implies that f is **not** locally invertible even though we have $f'(0) = 1 \neq 0$. One cannot apply the inverse function theorem in this case since f is not $\mathscr{C}^1$.

Corollary A.21. *Let $U \subseteq \mathbb{R}^n$ be open, $f : U \to \mathbb{R}^n$ be of class $\mathscr{C}^1$, and $(Df)(x)$ be invertible for all $x \in U$. Then $f(W)$ is open for every open set $W \subseteq U$.*

在證明了小區域的（local）反函數定理 (Theorem A.17) 之後，我們接下來要問的是全域的（global）反函數在什麼條件之下會存在。如果照一維的反函數定理，我們會猜測是不是只要 $(Df)(x)$ 在整個區域都可逆就能得到在全域的反函數都存在。以下給個反例說單單在這個條件之下，函數不一定會有一對一的性質。

Example A.22. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$f(x, y) = (e^x \cos y, e^x \sin y).$$

Then

$$\big[(Df)(x, y)\big] = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix}.$$

It is easy to see that the Jacobian of f at any point is not zero (thus $(Df)(x)$ is invertible for all $x \in \mathbb{R}^2$), and f is not globally one-to-one (thus the inverse of f does not exist globally) since for example, $f(x, y) = f(x, y + 2\pi)$.

要再加什麼條件進來才能得到反函數在全域都存在是個不容易的問題。在一維的情況下，導數是 sign definite 就表示函數在全域是嚴格單調的。在高維度的情況，即使是 $(Df)(x)$ 到處都可逆，仍然有很多情況可能發生（如上例）。下面這個定理（全域的反函數存在定理），從某種角度來說並沒有真的加了什麼條件以確保全域的反函數存在，只是多要求了在所考慮的區域邊界上函數是一對一的。這個條件在一維的情況之下是自動成立的：因為如果一單變數函數的導數是 sign definite，那麼函數在邊界上必定是一對一的（因為嚴格單調的關係）。

Theorem A.23 (Global Existence of Inverse Function)**.** *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open, $f : \mathcal{D} \to \mathbb{R}^n$ be of class $\mathscr{C}^1$, and $(Df)(x)$ be invertible for all $x \in \mathcal{D}$. Suppose that K is a connected compact subset of $\mathcal{D}$, and $f : \partial K \to \partial f(K)$ is one-to-one. Then $f : K \to \mathbb{R}^n$ is one-to-one.*

Proof. Define $E = \big\{x \in K \,\big|\, \exists\, y \in K, y \neq x \ni f(x) = f(y)\big\}$. Our goal is to show that $E = \varnothing$.

Before proceeding, we note that Corollary A.21 implies that if $x \in \text{int}(K)$, then $f(x) \in \text{int}(f(K))$. Therefore, $f : \partial K \to \partial f(K)$ and $f : \text{int}(K) \to \text{int}(f(K))$.

Claim 1: E is closed.

Proof of claim 1: Suppose the contrary that E is not closed. Then there exists $\{x_k\}_{k=1}^{\infty} \subseteq E$, $x_k \to x$ as $k \to \infty$ but $x \in K \backslash E$. Since $x_k \in E$, by the definition of E there exists $y_k \in E$ such that $y_k \neq x_k$ and $f(x_k) = f(y_k)$. By the compactness of K, there exists a convergent subsequence $\big\{y_{k_j}\big\}_{j=1}^{\infty}$ of $\{y_k\}_{k=1}^{\infty}$ with limit $y \in K$. Since $x \notin E$ and $f(x_{k_j}) = f(y_{k_j}) \to f(y)$ as $j \to \infty$, we must have $x = y$; thus $y_{k_j} \to x$ as $j \to \infty$.

Since $(Df)(x)$ is invertible, by the inverse function theorem there exists $\delta > 0$ such that $f : B(x,\delta) \to \mathbb{R}^n$ is one-to-one. By the convergence of sequences $\{x_{k_j}\}_{j=1}^{\infty}$ and $\{y_{k_j}\}_{j=1}^{\infty}$, there exists $N > 0$ such that

$$x_{k_j}, y_{k_j} \in B(x,\delta) \qquad \forall\, j \geqslant N\,.$$

This implies that $f : B(x,\delta) \to \mathbb{R}^n$ cannot be one-to-one $\big($since $x_{k_j} \neq y_{k_j}$ but $f\big(x_{k_j}\big) = f\big(y_{k_j}\big)\big)$, a contradiction. Therefore, E is closed.

Claim 2: E is open relative to K; that is, for every $x \in E$, there exists an open set U such that $x \in U$ and $U \cap K \subseteq E$.

Proof of claim 2: Let $x_1 \in E$. Then there is $x_2 \in E$, $x_2 \neq x_1$, such that $f(x_1) = f(x_2)$. By the assumption that f is injective from ∂K to the boundary of $f(K)$, we find that $x_1, x_2 \in \operatorname{int}(K)$. Since $(Df)(x_1)$ and $(Df)(x_2)$ are invertible, by the inverse function theorem there exist open neighborhoods U_1 of x_1 and U_2 of x_2, as well as open neighborhoods V_1, V_2 of $f(x_1)$, such that $f : U_1 \to V_1$ and $f : U_2 \to V_2$ are both one-to-one and onto. Since $x_1 \neq x_2$ and $x_1, x_2 \in \operatorname{int}(K)$, W.L.O.G. we can assume that $U_1 \cap U_2 = \varnothing$, $U_1, U_2 \subseteq K$ and $V_1, V_2 \subseteq f(K)$. Since $V_1 \cap V_2$ is open, the continuity of f implies that $f^{-1}(V_1 \cap V_2) = \mathcal{O} \cap \mathcal{D}$ for some open set $\mathcal{O}$; thus

$$\begin{aligned} f : U_1 \cap \mathcal{O} \to V_1 \cap V_2 \text{ is one-to-one and onto}\,,\\ f : U_2 \cap \mathcal{O} \to V_1 \cap V_2 \text{ is one-to-one and onto}\,. \end{aligned}$$

Let $U = U_1 \cap \mathcal{O}$. Then U is an open neighborhood of x_1. Moreover, every $x \in U$ corresponds to a unique $\widetilde{x} \in U_2 \cap \mathcal{O}$ such that $f(x) = f(\widetilde{x})$. Since $U_1 \cap U_2 = \varnothing$, we must have $x \neq \widetilde{x}$. Therefore, $x \in E$, or equivalently, $U \subseteq E$.

Now we show that $E = \varnothing$. Since E is open relative to K and E is closed, the connectedness of K implies that $E = K$ or $E = \varnothing$. Since $f : \partial K \to \partial f(K)$ and $f : \operatorname{int}(K) \to \operatorname{int}(f(K))$, the injectivity of f on ∂K implies that $\partial K \nsubseteq E$; thus $K \nsubseteq E$. Therefore, $E = \varnothing$. □

Example A.24. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be given as in Example A.22, and $\mathcal{D} = \big\{(x,y) \,\big|\, x \in \mathbb{R}, 0 < y < 2\pi\big\}$. Then $f : \mathcal{D} \to \mathbb{R}^2$ is one-to-one. If K is a compact subset of $\mathcal{D}$, then $f : K \to \mathbb{R}^2$ is also one-to-one (thus $f : \partial K \to \mathbb{R}^2$ must be one-to-one as well).

Corollary A.25. *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be a bounded open convex set, and $f : \mathcal{D} \to \mathbb{R}^n$ be of class $\mathscr{C}^1$ such that*

1. *f and Df are continuous on $\bar{\mathcal{D}}$;*

2. *the Jacobian* $\det\big([(Df)(x)]\big) \neq 0$ *for all $x \in \bar{\mathcal{D}}$;*

3. *$f : \partial\mathcal{D} \to \mathbb{R}^n$ is one-to-one.*

Then $f : \bar{\mathcal{D}} \to \mathbb{R}^n$ is one-to-one. Moreover, $f^{-1} : f(\bar{\mathcal{D}}) \to \mathbb{R}^n$ is continuous, and $f^{-1} : f(\mathcal{D}) \to \mathcal{D}$ is of class $\mathscr{C}^1$.

Proof. We first claim that there exists a small $\varepsilon > 0$ such that $f : \mathcal{D}_\varepsilon \to \mathbb{R}^n$ is one-to-one, where

$$\mathcal{D}_\varepsilon \equiv \big\{x \in \mathcal{D} \,\big|\, d(x, \partial\mathcal{D}) < \varepsilon\big\}.$$

Assume the contrary that for every $k > 0$, there exists $x_k, y_k \in \mathcal{D}$ such that

$$\text{(a)}\;\; x_k \neq y_k; \qquad \text{(b)}\; d(x_k, \partial\Omega) < \frac{1}{k} \text{ and } d(y_k, \partial\Omega) < \frac{1}{k}; \qquad \text{(c)}\; f(x_k) = f(y_k).$$

Since $\{x_k\}_{k=1}^\infty$ and $\{y_k\}_{k=1}^\infty$ are bounded (due to the boundedness of $\mathcal{D}$), by the Bolzano-Weierstrass Theorem (Theorem 2.51) there exist $\big\{x_{k_j}\big\}_{j=1}^\infty$ and $\big\{y_{k_j}\big\}_{j=1}^\infty$ such that $x_{k_j} \to x \in \bar{\mathcal{D}}$ and $y_{k_j} \to y \in \bar{\mathcal{D}}$. By (b), $x, y \in \partial\mathcal{D}$; thus the fact that $f : \partial\mathcal{D} \to \mathbb{R}^n$ is one-to-one implies that $x = y$. Therefore, $x_{k_j} \to x$ and $y_{k_j} \to x$ as $j \to \infty$.

Let $u_j = \dfrac{x_{k_j} - y_{k_j}}{\|x_{k_j} - y_{k_j}\|_{\mathbb{R}^n}}$. Since $\{u_j\}_{j=1}^\infty$ is bounded in $\mathbb{R}^n$, by the Bolzano-Weierstrass Theorem again there is a convergent subsequence $\big\{u_{j_\ell}\big\}_{\ell=1}^\infty$ with limit $u \neq 0$. Moreover, by the convexity of $\mathcal{D}$, the mean value theorem implies that for each $i = 1, \cdots, n$, there exists $c_{i\ell}$ on the line segment joining $x_{k_{j_\ell}}$ and $y_{k_{j_\ell}}$ such that

$$0 = f_i(x_{k_{j_\ell}}) - f_i(y_{k_{j_\ell}}) = (Df_i)(c_{i\ell})(x_{k_{j_\ell}} - y_{k_{j_\ell}}) = \|x_{k_{j_\ell}} - y_{k_{j_\ell}}\|_{\mathbb{R}^n}(Df_i)(c_{i\ell})\big(u_{j_\ell}\big)$$

which by (a) further shows that $(Df_i)(c_{i\ell})\big(u_{j_\ell}\big) = 0$ for all $i = 1, \cdots, n$ and $\ell \in \mathbb{N}$. Since $c_{i\ell} \to x$ as $\ell \to \infty$, passing $\ell \to \infty$ we conclude that $(Df_i)(x)(u) = 0$. This holds for each $i = 1, \cdots, n$; thus $(Df)(x)(u) = 0$. Therefore, $\det\big([(Df)(x)]\big) = 0$, a contradiction.

Now suppose that there exists $x, y \in \mathcal{D}$ such that $f(x) = f(y)$. Choose a compact set $K \subseteq \mathcal{D}$ such that $x, y \in K$ and $\partial K \subseteq \mathcal{D}_\varepsilon$ (this can be done, for example, by choosing that $K = \bar{\mathcal{D}} \backslash \mathcal{D}_\delta$ for some small $\delta > 0$). Since $f : \mathcal{D}_\varepsilon \to \mathbb{R}^n$ is one-to-one, $f : \partial K \to \mathbb{R}^n$ is one-to-one. By Theorem A.23, $f : K \to \mathbb{R}^n$ is one-to-one. Then $x = y$; thus $f : \mathcal{D} \to \mathbb{R}^n$ is one-to-one.

Next, we show that $f : \bar{\mathcal{D}} \to \mathbb{R}^n$ is one-to-one. Assume the contrary that there exists $x \in \mathcal{D}$ and $y \in \partial\mathcal{D}$ such that $f(x) = f(y)$. By the inverse function theorem there exists open neighborhood U of x and V of $f(x)$ such that $f : U \to V$ is one-to-one and onto. By choosing U even smaller if necessary, we can assume that there exists $\{y_k\}_{k=1}^{\infty} \subseteq \mathcal{D}\backslash U$ and $y_k \to y$ as $k \to \infty$. By the continuity of f, $f(y_k) \to f(y)$ as $k \to \infty$. However, since $f : \mathcal{D} \to \mathbb{R}^n$ is one-to-one, $\big\{f(y_k)\big\}_{k=1}^{\infty} \notin V$; thus $\big\{f(y_k)\big\}_{k=1}^{\infty}$ cannot converge to $f(y)$ as $k \to \infty$ (since $f(y) \in V$), a contradiction.

Finally, the inverse function theorem implies that $f^{-1} : f(\mathcal{D}) \to \mathcal{D}$ is of class $\mathscr{C}^1$, and the continuity of f^{-1} on $f(\bar{\mathcal{D}})$ follows from the fact that $(f^{-1})^{-1}(F) = f(F)$ is closed in $\bar{\mathcal{D}}$ for all closed subset F of $\mathbb{R}^n$. □

Remark A.26. Suppose that $\mathcal{D} \subseteq \mathbb{R}^n$ in Corollary A.25 is open, bounded, connected but **not** convex. The Whitney extension theorem (which is not covered in this text) implies that there exists a function $F : \mathbb{R}^n \to \mathbb{R}^n$ so that $F = f$ and $DF = Df$ on $\bar{\mathcal{D}}$. Then Theorem A. 23 can be applied to guarantee that F is one-to-one on $\bar{\mathcal{D}}$.

A.4 The Implicit Function Theorem（隱函數定理）

Theorem A.27 (Implicit Function Theorem). *Let $\mathcal{D} \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be open, and $F : \mathcal{D} \to \mathbb{R}^m$ be a function of class $\mathscr{C}^1$. Suppose that for some $(x_0, y_0) \in \mathcal{D}$, where $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^m$, $F(x_0, y_0) = 0$ and*

$$\big[(D_yF)(x_0,y_0)\big] = \begin{bmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_m} \end{bmatrix} (x_0, y_0)$$

is invertible. Then there exists an open neighborhood $U \subseteq \mathbb{R}^n$ of x_0, an open neighborhood $V \subseteq \mathbb{R}^m$ of y_0, and $f : U \to V$ such that

1. *$F\big(x, f(x)\big) = 0$ for all $x \in U$;*

2. *$y_0 = f(x_0)$;*

3. *$(Df)(x) = -\big((D_yF)(x, f(x))\big)^{-1}(D_xF)\big(x, f(x)\big)$ for all $x \in U$, where the matrix rep-*

resentation of $D_xF(x, f(x)) \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$ *is given by*

$$\big[(D_xF)(x,y)\big] = \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial x_1} & \cdots & \dfrac{\partial F_m}{\partial x_n} \end{bmatrix} (x,y).$$

4. f *is of class* $\mathscr{C}^1$*;*

5. *If* F *is of class* $\mathscr{C}^r$ *for some* $r > 1$*, so is* f*.*

Proof. Let $z = (x,y)$ and $w = (u,v)$, where $x,u \in \mathbb{R}^n$ and $y,v \in \mathbb{R}^m$. Define G by $G(x,y) = \big(x, F(x,y)\big)$, and write $w = G(z)$. Then $G : \mathcal{D} \to \mathbb{R}^{n+m}$, and

$$\big[(DG)(x,y)\big] = \begin{bmatrix} \mathbb{I}_n & 0 \\ (D_xF)(x,y) & (D_yF)(x,y) \end{bmatrix},$$

where $\mathbb{I}_n$ is the $n \times n$ identity matrix. We note that the Jacobian of G at (x_0, y_0) is $\det\big([(D_yF)(x_0,y_0)]\big)$ which does not vanish since $(D_yF)(x_0,y_0)$ is invertible, so the inverse function theorem implies that there exists open neighborhoods $\mathcal{O}$ of (x_0,y_0) and $\mathcal{W}$ of $\big(x_0, F(x_0,y_0)\big) = (x_0, 0)$ such that

(a) $G : \mathcal{O} \to \mathcal{W}$ is one-to-one and onto;

(b) the inverse function $G^{-1} : \mathcal{W} \to \mathcal{O}$ is of class $\mathscr{C}^r$;

(c) $(DG^{-1})\big(x, F(x,y)\big) = \big((DG)(x,y)\big)^{-1}$.

By Remark A.18, W.L.O.G. we can assume that $\mathcal{O} = U \times V$, where $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ are open, and $x_0 \in U$, $y_0 \in V$.

Write $G^{-1}(u,v) = \big(\varphi(u,v), \psi(u,v)\big)$, where $\varphi : \mathcal{W} \to U$ and $\psi : \mathcal{W} \to V$. Then

$$(u,v) = G\big(\varphi(u,v), \psi(u,v)\big) = \big(\varphi(u,v), F(u, \psi(u,v))\big)$$

which implies that $\varphi(u,v) = u$ and $v = F(u, \psi(u,v))$. Let $f(x) = \psi(x, 0)$. Then $\big(u, f(u)\big) \in U \times V$ is the unique point satisfying $F\big(u, f(u)\big) = 0$ if $u \in U$. Therefore, $f : U \to V$, and

$$F\big(x, f(x)\big) = 0 \qquad \forall\, x \in U.$$

Since $G(x_0, y_0) = (x_0, 0) = G\big(x_0, f(x_0)\big)$, $(x_0, y_0), \big(x_0, f(x_0)\big) \in \mathcal{O}$, and $G : \mathcal{O} \to \mathcal{W}$ is one-to-one, we must have $y_0 = f(x_0)$.

By (b) and (c), we have G^{-1} is of class $\mathscr{C}^1$, and

$$(DG^{-1})(u, v) = \big((DG)(x, y)\big)^{-1}.$$

As a consequence, $\psi \in \mathscr{C}^1$, and

$$\begin{aligned}\begin{bmatrix}(D_u\varphi)(u,v) & (D_v\varphi)(u,v)\\ (D_u\psi)(u,v) & (D_v\psi)(u,v)\end{bmatrix} &= \begin{bmatrix}\mathbb{I}_n & 0\\ (D_xF)(x,y) & (D_yF)(x,y)\end{bmatrix}^{-1}\\ &= \begin{bmatrix}\mathbb{I}_n & 0\\ -\big((D_yF)(x,y)\big)^{-1}(D_xF)(x,y) & \big((D_yF)(x,y)\big)^{-1}\end{bmatrix}.\end{aligned}$$

Evaluating the equation above at $v = 0$, we conclude that

$$(Df)(u) = (D_u\psi)(u, 0) = -\big((D_yF)(u, f(u))\big)^{-1}(D_xF)\big(u, f(u)\big)$$

which implies 3. We also note that 4 follows from (b) and 5 follows from 3. $\square$

Alternative proof of Theorem A.27 ***without*** *applying the inverse function theorem.* Let $z = (x, y)$, $z_0 = (x_0, y_0)$, $A = (D_xF)(x_0, y_0)$ and $B = (D_yF)(x_0, y_0)$. Define

$$r(x, y) = F(x, y) - A(x - x_0) - B(y - y_0).$$

Note that the condition $F(x_0, y_0) = 0$ implies that $r(x_0, y_0) = 0$. Our goal is to solve the equation

$$0 = A(x - x_0) + B(y - y_0) + r(x, y)$$

for y (given x). By the invertibility of B, this is equivalent of finding a fixed-point of the map

$$\Phi_x(y) = y_0 - B^{-1}\big[A(x - x_0) + r(x, y)\big].$$

Since r is of class $\mathscr{C}^1$ and $(Dr)(x_0, y_0) = 0$, there exists $\delta > 0$ such that

$$\|(Dr)(x, y)\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)} < \min\Big\{\frac{1}{4m\|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}}, \frac{1}{2m}\Big\} \qquad \forall\, z \in B(z_0, \delta). \tag{A.4.1}$$

Therefore, the mean value theorem (Theorem A.16) implies that

$$\begin{aligned}\|r(x, y)\|_{\mathbb{R}^m} &= \big\|r(x, y) - r(x_0, y_0)\big\|_{\mathbb{R}^m} \leqslant \sum_{i=1}^m \big|r_i(x, y) - r_i(x_0, y_0)\big| = \sum_{i=1}^m \big|(Dr_i)(c_i)(z - z_0)\big|\\ &\leqslant \frac{\|z - z_0\|_{\mathbb{R}^{n+m}}}{4\|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}} < \frac{\delta}{4\|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}}\end{aligned}$$

for all $z = (x,y) \in B\big(x_0, \dfrac{\delta}{2}\big) \times B\big(y_0, \dfrac{\delta}{2}\big) \subseteq B(z_0, \delta)$, and

$$\big\|\Phi_x(y_1) - \Phi_x(y_2)\big\|_{\mathbb{R}^m} \leqslant \|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}\|r(x,y_1) - r(x,y_2)\|_{\mathbb{R}^m} < \frac{1}{4}\|y_1 - y_2\|_{\mathbb{R}^m} \tag{A.4.2}$$

if $x \in B\big(x_0, \dfrac{\delta}{2}\big)$, $y_1, y_2 \in B\big(y_0, \dfrac{\delta}{2}\big)$ and $y_1 \neq y_2$. As a consequence, for each (fixed) x satisfying $\|x - x_0\|_{\mathbb{R}^n} < r \equiv \min\Big\{\dfrac{\delta}{4\big(1 + \|A\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}\big)\|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}}, \dfrac{\delta}{2}\Big\}$, if $\|y - y_0\|_{\mathbb{R}^m} < \dfrac{\delta}{2}$ we have

$$\begin{aligned}\|\Phi_x(y) - y_0\|_{\mathbb{R}^n} &\leqslant \|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}\big\|A(x - x_0) + r(x,y)\big\|_{\mathbb{R}^m} \\ &\leqslant \|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}\big[\|A\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}\|x - x_0\|_{\mathbb{R}^n} + \|r(x,y)\|_{\mathbb{R}^m}\big] < \frac{\delta}{2}.\end{aligned} \tag{A.4.3}$$

Let $M = \big\{y \in \mathbb{R}^m \,\big|\, \|y - y_0\|_{\mathbb{R}^m} \leqslant \dfrac{\delta}{2}\big\}$. Then for each $x \in U \equiv B(x_0, r)$, (A.4.2) and (A.4.3) imply that $\Phi_x : M \to M$ is a contraction mapping; thus there is a unique fixed-point $y \in M$. Denote this unique fixed-point as $f(x)$. Then $f : U \to V \equiv B\big(y_0, \dfrac{\delta}{2}\big)$ (the choice of this V guarantees that $U \times V \subseteq B(z_0, \delta)$) and

$$F\big(x, f(x)\big) = A(x - x_0) + B\big(f(x) - y_0\big) + r\big(x, f(x)\big) = 0\,.$$

Moreover, since $F(x,y) = 0$ if and only if y is a fixed-point of Φ_x, and the contraction mapping principle provides the uniqueness of the fixed-point if $(x,y) \in U \times M$. Since $(x_0,y_0), (x_0, f(x_0)) \in U \times M$, we must have $y_0 = f(x_0)$.

To see the differentiability of f, we first claim that $f : U \to V$ is continuous. Since $f(x)$ is the fixed-point of Φ_x,

$$f(x) = y_0 - B^{-1}\big(A(x - x_0) + r(x, f(x))\big)\,.$$

If $x_1, x_2 \in U$, then $(x_1, f(x_1)), (x_2, f(x_2)) \in B(z_0, \delta)$; thus (A.4.1) (together with the mean value theorem) (A.4.2) imply that

$$\begin{aligned}\big\|f(x_1) - f(x_2)\big\|_{\mathbb{R}^m} &= \big\|B^{-1}A(x_1 - x_2)\big\|_{\mathbb{R}^m} + \big\|r(x_1, f(x_1)) - r(x_2, f(x_2))\big\|_{\mathbb{R}^m} \\ &\leqslant \|B^{-1}A\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}\|x_1 - x_2\|_{\mathbb{R}^n} \\ &\quad + \|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)} \sup_{(x,y)\in U\times V}\big\|Dr(x,y)\big\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}\big\|(x_1, f(x_1)) - (x_2, f(x_2))\big\|_{\mathbb{R}^m} \\ &\leqslant \|B^{-1}A\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}\|x_1 - x_2\|_{\mathbb{R}^n} + \frac{1}{2}\|x_1 - x_2\|_{\mathbb{R}^n} + \frac{1}{2}\|f(x_1) - f(x_2)\|_{\mathbb{R}^m}\,.\end{aligned}$$

Therefore,

$$\big\|f(x_1) - f(x_2)\big\|_{\mathbb{R}^m} \leqslant \big(2\|B^{-1}A\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} + 1\big)\|x_1 - x_2\|_{\mathbb{R}^n} \tag{A.4.4}$$

which implies that $f : U \to V$ is (Lipschitz) continuous.

Now let $a \in U$ and $\varepsilon > 0$ be given. Define $b = (a, f(a))$, and $\widetilde{A} = (D_x F)(b)$, $\widetilde{B} = (D_y F)(b)$. We would like to show that there exists $\delta_1 > 0$ such that

$$\big\|f(x) - f(a) + \widetilde{B}^{-1}\widetilde{A}(x-a)\big\|_{\mathbb{R}^m} \leqslant \varepsilon \|x-a\|_{\mathbb{R}^n} \qquad \forall\, x \in B(a,\delta_1)\,.$$

Since $F \in \mathscr{C}^1$ and the map $L \mapsto L^{-1}$ is continuous, there exists $\delta_2 > 0$ such that

$$\big\|(D_y F)(z)^{-1}(D_x F)(z) - (D_y F)(z_0)^{-1}(D_x F)(z_0)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)} \leqslant \frac{\varepsilon}{4} \qquad \forall\, z \in B(z_0,\delta_2)\,.$$

Moreover, since $r \in \mathscr{C}^1$, there exists $\delta_3 > 0$ such that

$$\big\|r(z) - r(b) - (Dr)(b)(z-b)\big\|_{\mathbb{R}^m} \leqslant \frac{\varepsilon}{2\|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}}\|z-b\|_{\mathbb{R}^{n+m}} \quad \forall\, z \in B(b,\delta_3)$$

and

$$\big\|(Dr)(z) - (Dr)(b)\big\|_{\mathscr{B}(\mathbb{R}^{n+m},\mathbb{R}^m)} \leqslant \frac{\varepsilon}{2\|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}}\|z-b\|_{\mathbb{R}^{n+m}} \quad \forall\, z \in B(b,\delta_3)\,.$$

Choose $\delta_1 = \min\Big\{\dfrac{\delta_2}{2}, \dfrac{\delta_3}{2}, \dfrac{\delta_3}{2(2\|B^{-1}A\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}+1)}\Big\}$. Then if $\|x-a\|_{\mathbb{R}^n} < \delta_1$, using (A.4.4) we find that

$$\big\|(x,f(x)) - (a,f(a))\big\|_{\mathbb{R}^{n+m}} \leqslant \|x-a\|_{\mathbb{R}^n} + \|f(x)-f(a)\|_{\mathbb{R}^m} < \min\{\delta_2,\delta_3\}\,;$$

thus if $\|x-a\|_{\mathbb{R}^n} < \delta_1$,

$$\begin{aligned}
&\big\|f(x) - f(a) + \widetilde{B}^{-1}\widetilde{A}(x-a)\big\|_{\mathbb{R}^m} \\
&\qquad = \big\|\big(\widetilde{B}^{-1}\widetilde{A} - B^{-1}A\big)(x-a) + B^{-1}\big(r(x,f(x)) - r(a,f(a))\big)\big\|_{\mathbb{R}^m} \\
&\qquad \leqslant \big\|\widetilde{B}^{-1}\widetilde{A} - B^{-1}A\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}\|x-a\|_{\mathbb{R}^n} \\
&\qquad\quad + \|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}\big\|r(x,f(x)) - r(a,f(a)) - (Dr)(b)\big(x-a, f(x)-f(a)\big)\big\|_{\mathbb{R}^m} \\
&\qquad\quad + \|B^{-1}\|_{\mathscr{B}(\mathbb{R}^m,\mathbb{R}^m)}\big\|(Dr)(b)\big((x-a, f(x)-f(a)\big)\big\|_{\mathbb{R}^m} \leqslant \varepsilon\|x-a\|_{\mathbb{R}^n}\,.
\end{aligned}$$

Therefore, f is differentiable on U, and

$$(Df)(x) = -\big((D_y F)(x,f(x))\big)^{-1}(D_x F)\big(x,f(x)\big) \qquad \forall\, x \in U\,. \tag{A.4.5}$$

Since F is of class $\mathscr{C}^1$ and f is continuous on U, we find that Df is continuous; thus f is of class $\mathscr{C}^1$. $\square$

Example A.28. Let $F(x,y) = x^2 + y^2 - 1$.

1. If $(x_0, y_0) = (1, 0)$, then $F_x(x_0, y_0) = 2 \neq 0$; thus the implicit function theorem implies that locally x can be expressed as a function of y.

2. If $(x_0, y_0) = (0, -1)$, then $F_y(x_0, y_0) = -2 \neq 0$; thus the implicit function theorem implies that locally y can be expressed as a function of x.

3. If $(x_0, y_0) = \big(-\frac{1}{2}, \frac{\sqrt{3}}{2}\big)$, then $F_x(x_0, y_0) = -1 \neq 0$ and $F_y(x_0, y_0) = \sqrt{3} \neq 0$; thus the implicit function theorem implies that locally x can be expressed as a function of y and locally y can be expressed as a function of x.

Example A.29. Suppose that (x, y, u, v) satisfies the equation

$$\left\{\begin{aligned} xu + yv^2 &= 0 \\ xv^3 + y^2u^6 &= 0 \end{aligned}\right.$$

and $(x_0, y_0, u_0, v_0) = (1, -1, 1, -1)$. Let $F(x, y, u, v) = (xu + yv^2, xv^3 + y^2u^6)$. Then $F(x_0, y_0, u_0, v_0) = 0$.

1. Since $(D_{x,y}F)(x_0, y_0, u_0, v_0) = \begin{bmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} \end{bmatrix}(x_0, y_0, u_0, v_0) = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ is invertible, locally (x, y) can be expressed in terms of u, v; that is, locally $x = x(u, v)$ and $y = y(u, v)$.

2. Since $(D_{y,u}F)(x_0, y_0, u_0, v_0) = \begin{bmatrix} \dfrac{\partial F_1}{\partial y} & \dfrac{\partial F_1}{\partial u} \\ \dfrac{\partial F_2}{\partial y} & \dfrac{\partial F_2}{\partial u} \end{bmatrix}(x_0, y_0, u_0, v_0) = \begin{bmatrix} 1 & 1 \\ -2 & 6 \end{bmatrix}$ is invertible, locally (y, u) can be expressed in terms of x, v.

Example A.30. Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$f(x, y, z) = (xe^y + ye^z, xe^z + ze^y).$$

Then f is of class $\mathscr{C}^1$, $f(-1, 1, 1) = (0, 0)$ and

$$\big[(Df)(x, y, z)\big] = \begin{bmatrix} e^y & xe^y + e^z & ye^z \\ e^z & ze^y & xe^z + e^y \end{bmatrix}.$$

Since $(D_{y,z}f)(-1,1,1) = \begin{bmatrix} 0 & e \\ e & 0 \end{bmatrix}$ is invertible, the implicit function theorem implies that the system

$$\begin{cases} xe^y + ye^z = 0 \\ xe^z + ze^y = 0 \end{cases}$$

can be solved for y and z as continuously differentiable function of x for x near -1 and (y,z) near $(1,1)$. Furthermore, if we write $(y,z) = g(x)$ for x near -1, then

$$g'(x) = \begin{bmatrix} xe^y + e^z & ye^z \\ ze^y & xe^z + e^y \end{bmatrix}^{-1} \begin{bmatrix} e^y \\ e^z \end{bmatrix}.$$

A.5 An Application: The Lagrange Multipliers

In this section we are concerned with the optimization problem

$$\text{“find the extreme value of function } y = f(x) \text{ subject to the constraint } g(x) = 0\text{”}, \quad \text{(A.5.1)}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a real-value differentiable function, and $g : \mathbb{R}^n \to \mathbb{R}^m$ is a vector-valued $\mathscr{C}^1$-function for some $m < n$. We assume that the feasible set $\{x \,|\, g(x) = 0\}$ is non-empty and does not contain isolated points (or the extreme value of f is not attained at isolated points of the feasible set), and all the local extreme points of f do not belong to the feasible set (otherwise we can ignore the constraint and find the local extreme value of f directly).

Suppose that f attains its extreme value, subject to the constraint $g = 0$, at point a, and the Jacobian matrix of g at a has rank m (that is, $[(Dg)(a)]$ has full rank). Without loss of generality we can assume that the square matrix

$$\begin{bmatrix} \dfrac{\partial g_1}{\partial x_1}(a) & \cdots & \dfrac{\partial g_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(a) & \cdots & \dfrac{\partial g_m}{\partial x_m}(a) \end{bmatrix}$$

is invertible (thus has rank m). Then the implicit function theorem implies there exist an open neighborhood V of $(a_{m+1}, \cdots, a_n)$ and a $\mathscr{C}^1$-function $\varphi : V \to \mathbb{R}^m$ such that in a neighborhood of a, the feasible set can be expressed as

$$(x_1, \cdots, x_m) = \varphi(x_{m+1}, \cdots, x_n) \qquad (x_{m+1}, \cdots, x_n) \in V\,.$$

Under these settings, the original optimization problem (A.5.1) is transformed as "the function $y = f\big(\varphi(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n\big)$ attains its extreme value at $(a_{m+1}, \cdots, a_n)$"; thus

$$\begin{aligned}
\frac{\partial}{\partial x_{m+1}}\Big|_{(x_{m+1},\cdots,x_n)=(a_{m+1},\cdots,a_n)} f\big(\varphi(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n\big) &= 0\,,\\
\frac{\partial}{\partial x_{m+2}}\Big|_{(x_{m+1},\cdots,x_n)=(a_{m+1},\cdots,a_n)} f\big(\varphi(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n\big) &= 0\,,\\
&\vdots\\
\frac{\partial}{\partial x_{n}}\Big|_{(x_{m+1},\cdots,x_n)=(a_{m+1},\cdots,a_n)} f\big(\varphi(x_{m+1}, \cdots, x_n), x_{m+1}, \cdots, x_n\big) &= 0\,,
\end{aligned}$$

or using the chain rule,

$$\begin{bmatrix} \dfrac{\partial f}{\partial x_1}(a) & \cdots & \dfrac{\partial f}{\partial x_m}(a) \end{bmatrix} \frac{\partial \varphi}{\partial x_j}(a_{m+1}, \cdots, a_n) + \frac{\partial f}{\partial x_j}(a) = 0 \quad \text{for } m+1 \leqslant j \leqslant n\,. \tag{A.5.2}$$

Noting that the implicit function theorem implies that

$$\big[(D\varphi)(a_{m+1}, \cdots, a_n)\big] = -\begin{bmatrix} \dfrac{\partial g_1}{\partial x_1}(a) & \cdots & \dfrac{\partial g_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(a) & \cdots & \dfrac{\partial g_m}{\partial x_m}(a) \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{\partial g_1}{\partial x_{m+1}}(a) & \cdots & \dfrac{\partial g_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \dfrac{\partial g_m}{\partial x_{m+1}}(a) & \cdots & \dfrac{\partial g_m}{\partial x_n}(a) \end{bmatrix},$$

we find that for $m+1 \leqslant j \leqslant n$,

$$\frac{\partial \varphi}{\partial x_j}(a_{m+1}, \cdots, a_n) = -\begin{bmatrix} \dfrac{\partial g_1}{\partial x_1}(a) & \cdots & \dfrac{\partial g_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(a) & \cdots & \dfrac{\partial g_m}{\partial x_m}(a) \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{\partial g_1}{\partial x_j}(a) \\ \vdots \\ \dfrac{\partial g_m}{\partial x_j}(a) \end{bmatrix}.$$

Define

$$\begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_m \end{bmatrix} = -\begin{bmatrix} \dfrac{\partial f}{\partial x_1}(a) & \cdots & \dfrac{\partial f}{\partial x_m}(a) \end{bmatrix} \begin{bmatrix} \dfrac{\partial g_1}{\partial x_1}(a) & \cdots & \dfrac{\partial g_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(a) & \cdots & \dfrac{\partial g_m}{\partial x_m}(a) \end{bmatrix}^{-1}.$$

Then (A.5.2) implies that

$$\frac{\partial f}{\partial x_j}(a) + \lambda_1 \frac{\partial g_1}{\partial x_j}(a) + \lambda_2 \frac{\partial g_2}{\partial x_j}(a) + \cdots + \lambda_m \frac{\partial g_m}{\partial x_j}(a) = 0 \qquad \text{for } m+1 \leqslant j \leqslant n. \tag{A.5.3}$$

On the other hand, by the fact that

$$\begin{bmatrix} \dfrac{\partial g_1}{\partial x_1}(a) & \cdots & \dfrac{\partial g_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(a) & \cdots & \dfrac{\partial g_m}{\partial x_m}(a) \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{\partial g_1}{\partial x_1}(a) & \cdots & \dfrac{\partial g_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(a) & \cdots & \dfrac{\partial g_m}{\partial x_m}(a) \end{bmatrix} = \mathrm{I}_{m\times m}\,,$$

for $1 \leqslant j \leqslant m$,

$$\begin{bmatrix} \dfrac{\partial g_1}{\partial x_1}(a) & \cdots & \dfrac{\partial g_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(a) & \cdots & \dfrac{\partial g_m}{\partial x_m}(a) \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{\partial g_1}{\partial x_j}(a) \\ \vdots \\ \dfrac{\partial g_m}{\partial x_j}(a) \end{bmatrix} = \mathrm{e}_j\,,$$

where $\{\mathrm{e}_j\}_{j=1}^m$ is the standard basis of $\mathbb{R}^m$. Therefore, left multiplying the equation above by the row vector $-\begin{bmatrix} \dfrac{\partial f}{\partial x_1}(a) & \cdots & \dfrac{\partial f}{\partial x_m}(a) \end{bmatrix}$, we obtain that

$$\lambda_1 \frac{\partial g_1}{\partial x_j}(a) + \lambda_2 \frac{\partial g_2}{\partial x_j}(a) + \cdots + \lambda_m \frac{\partial g_m}{\partial x_j}(a) = -\frac{\partial f}{\partial x_j}(a) \qquad \text{for } 1 \leqslant j \leqslant m. \tag{A.5.4}$$

Combining (A.5.3) and (A.5.4), we conclude that

$$\frac{\partial f}{\partial x_j}(a) + \lambda_1 \frac{\partial g_1}{\partial x_j}(a) + \lambda_2 \frac{\partial g_2}{\partial x_j}(a) + \cdots + \lambda_m \frac{\partial g_m}{\partial x_j}(a) = 0 \qquad \text{for } 1 \leqslant j \leqslant n$$

or equivalently,

$$\big[(Df)(a)\big] + \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_m \end{bmatrix} \big[(Dg)(a)\big] = 0\,.$$

We then establish the following

Theorem A.31 (Lagrange Multipliers)**.** *Let $U \subseteq \mathbb{R}^n$ be an open set, $f : U \to \mathbb{R}$ be differentiable, $g : U \to \mathbb{R}^m$ be of class $\mathscr{C}^1$ for some $m < n$, and $\mathcal{S} = \big\{x \in U \,\big|\, g(x) = 0\big\}$. Suppose that $\mathcal{S} \neq \varnothing$ and $f|_{\mathcal{S}}$, the restriction of f on $\mathcal{S}$, attains its extreme value at $a \in \mathcal{S}$. If $\big[(Dg)(a)\big]$ has full rank, then there exist $\lambda_1, \lambda_2, \cdots, \lambda_m \in \mathbb{R}$ such that*

$$\big[(Df)(a)\big] + \begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_m \end{bmatrix} \big[(Dg)(a)\big] = 0\,.$$

A.6 Exercise

§ A.3 The Inverse Function Theorem

Problem A.1. Prove Corollary A.21; that is, show that if $U \subseteq \mathbb{R}^n$ is open, $f : U \to \mathbb{R}^n$ is of class $\mathscr{C}^1$, and $(Df)(x)$ is invertible for all $x \in U$, then $f(\mathcal{W})$ is open for every open set $\mathcal{W} \subseteq U$.

§ A.4 The Implicit Function Theorem

Problem A.2. Assume that one proves the implicit function theorem without applying the inverse theorem. Show the inverse function using the implicit function theorem.

Problem A.3. Suppose that $F(x, y, z) = 0$ is such that the functions $z = f(x, y)$, $x = g(y, z)$, and $y = h(z, x)$ all exist by the implicit function theorem. Show that $f_x \cdot g_y \cdot h_z = -1$.

Problem A.4. Suppose that the implicit function theorem applies to $F(x, y) = 0$ so that $y = f(x)$. Find a formula for f'' in terms of F and its partial derivatives. Similarly, suppose that the implicit function theorem applies to $F(x_1, x_2, y) = 0$ so that $y = f(x_1, x_2)$. Find formulas for $f_{x_1x_1}$, $f_{x_1x_2}$ and $f_{x_2x_2}$ in terms of F and its partial derivatives.

Problem A.5. Given $F : \mathbb{R}^4 \to \mathbb{R}^2$ and suppose that $F(x, y) = 0$, where $x = (x_1, x_2)$ and $y = (y_1, y_2)$. State conditions which guarantees that the equation

$$\frac{\partial y_1}{\partial x_1}\frac{\partial x_2}{\partial y_1} + \frac{\partial y_2}{\partial x_1}\frac{\partial x_2}{\partial y_2} = 0$$

holds. Prove or justify your answer.

§ A.5 An Application: The Lagranege Multipliers

Problem A.6. Find the extreme value of $f(x, y) = 4xy$ subject to the constraint

$$\frac{x^2}{9} + \frac{y^2}{16} = 1\,.$$

Problem A.7. Find the extreme value of $f(x, y) = x^2 + 6(y^2 + y + 1)^2$ subject to the constraint $x^2 + (y^3 - 1)^2 = 1$.

Problem A.8. Find the minimum value of $f(x, y, z) = 2x^2 + y^2 + 3z^2$ subject to the constraint $2x - 3y - 4z = 49$.

Problem A.9. Find the extreme value of the function $f(x, y, z) = 20 + 2x + 2y + z^2$ subject to two constraints $x^2 + y^2 + z^2 = 11$ and $x + y + z = 3$.

Problem A.10. Find the extreme value of $f(x,y,z) = z$ subject to the constraints $x^4 + y^4 - z^3 = 0$ and $y = z$.

Problem A.11. Suppose that $C \subseteq \mathbb{R}^n$ is a convex set (that is, $t\boldsymbol{x} + (1-t)\boldsymbol{y} \in C$ for all $\boldsymbol{x}, \boldsymbol{y} \in C$ and $t \in [0,1]$), and $f : C \to \mathbb{R}$ be a differentiable real-valued function. Show that if f on C attains its minimum at a point $\boldsymbol{x}^*$, then

$$(\nabla f)(\boldsymbol{x}^*) \cdot (\boldsymbol{x} - \boldsymbol{x}^*) \geqslant 0 \qquad \forall\, \boldsymbol{x} \in C. \tag{$\star$}$$

Hint: Recall that $(\nabla f)(\boldsymbol{x}^*) \cdot (\boldsymbol{x} - \boldsymbol{x}^*)$, when f is differentiable at x^*, is the directional derivative of f at $\boldsymbol{x}^*$ in the "direction" $(\boldsymbol{x} - \boldsymbol{x}^*)$.
Remark: A point $\boldsymbol{x}^*$ satisfying $(\star)$ is sometimes called a ***stationary point*** of f in C.

Problem A.12. Let B be the unit closed ball centered at the origin given by

$$B = \left\{\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n \,\middle|\, \|\boldsymbol{x}\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2 \leqslant 1\right\},$$

and $f : B \to \mathbb{R}$ be a differentiable real-valued function. Consider the minimization problem $\min\limits_{\boldsymbol{x} \in B} f(\boldsymbol{x})$.

(1) Show that if f attains its minimum at $\boldsymbol{x}^* \in B$, then there exists $\lambda \leqslant 0$ such that

$$(\nabla f)(\boldsymbol{x}^*) = \lambda \boldsymbol{x}^*.$$

(2) Find the minimum of the function $f(x,y) = x^2 + 2y^2 - x$ on the unit closed disk centered at the origin $\left\{(x,y) \,\middle|\, x^2 + y^2 \leqslant 1\right\}$ using (1).

Problem A.13. Let $\boldsymbol{a} \in \mathbb{R}^3$ be a vector, $b \in \mathbb{R}$, and C be a half plane given by

$$C = \left\{\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \,\middle|\, \boldsymbol{a} \cdot \boldsymbol{x} \leqslant b\right\},$$

and $f : C \to \mathbb{R}$ be a differentiable real-valued function. Consider the minimization problem $\min\limits_{\boldsymbol{x} \in C} f(\boldsymbol{x})$. Show that if f attains its minimum at $\boldsymbol{x}^* \in C$, then there exists $\lambda \leqslant 0$ such that

$$(\nabla f)(\boldsymbol{x}^*) = \lambda \boldsymbol{a}.$$

Appendix B

The Arzelà–Ascoli Theorem

在這一部份的附錄中，我們將研究一般情況下，連續函數列的逐點收斂與均勻收斂之間的具體差異為何。更具體地說，我們希望能找到一個條件，使得逐點收斂的連續函數列，其均勻收斂性等價於該條件成立。這個條件，刻劃了均勻收斂與逐點收斂的真實差異，而這個特別的條件，也將提供額外（且有效）的判斷法，幫助我們判斷在連續函數空間裡面的集合是否緊緻。

B.1 Equi-continuous family of functions

The first part of this section is devoted to the investigation of the difference between the pointwise convergence and the uniform convergence of sequence of continuous functions.

Definition B.1. Let (M, d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $A \subseteq M$ be a subset. A subset $B \subseteq \mathscr{C}_b(A;\mathcal{V})$ is said to be ***equi-continuous***（等度連續）if

$$\forall\, \varepsilon > 0, \exists\, \delta > 0 \ni \|f(x_1) - f(x_2)\| < \varepsilon \quad \text{whenever } d(x_1, x_2) < \delta,\ x_1, x_2 \in A, \text{ and } f \in B.$$

Remark B.2.
1. If $B \subseteq \mathscr{C}_b(A;\mathcal{V})$ is equi-continuous, and C is a subset of B, then C is also equi-continuous.
2. In an equi-continuous set of functions B, every $f \in B$ is uniformly continuous.

Remark B.3. For a uniformly continuous function f, let $\delta_f(\varepsilon)$ (we have defined this number in Remark 4.47) denote the largest δ that can be used in the definition of the uniform continuity; that is, $\delta_f(\varepsilon)$ has the property that

$$\|f(x) - f(y)\| < \varepsilon \text{ whenever } d(x, y) < \delta, x, y \in A \quad \Leftrightarrow \quad 0 < \delta \leqslant \delta_f(\varepsilon).$$

Suppose that every element in $B \subseteq \mathscr{C}_b(A;\mathcal{V})$ is uniformly continuous on A. Then B is equi-continuous if and only if $\inf\limits_{f\in B} \delta_f(\varepsilon) > 0$.

Example B.4. Let $B = \big\{f \in \mathscr{C}_b((0,1);\mathcal{V}) \,\big|\, |f'(x)| \leqslant 1 \text{ for all } x \in (0,1)\big\}$. Then B is equi-continuous (by choosing $\delta = \epsilon$ for any given ϵ, and applying the mean value theorem).

Example B.5. Let $f_k : [0,1] \to \mathbb{R}$ be a sequence of functions given by

$$f_k(x) = \begin{cases} kx & \text{if } 0 \leqslant x \leqslant \dfrac{1}{k}, \\ 2 - kx & \text{if } \dfrac{1}{k} \leqslant x \leqslant \dfrac{2}{k}, \\ 0 & \text{if } x \geqslant \dfrac{2}{k}, \end{cases}$$

and $B = \{f_k\}_{k=1}^{\infty}$. Then B is not equi-continuous since the largest δ for each k is $\dfrac{\varepsilon}{k}$ which converges to 0.

Lemma B.6. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $K \subseteq M$ be a compact subset. If $B \subseteq \mathscr{C}(K;\mathcal{V})$ is pre-compact, then B is equi-continuous.*

Proof. Suppose the contrary that B is not equi-continuous. Then there exists $\varepsilon > 0$ such that

$$\forall\, k \in \mathbb{N}, \exists\, x_k, y_k \in K \text{ and } f_k \in B \ni d(x_k, y_k) < \frac{1}{k} \text{ but } \|f_k(x_k) - f_k(y_k)\| \geqslant \varepsilon\,.$$

Since B is pre-compact in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$ and K is compact in (M,d), there exists a subsequence $\big\{f_{k_j}\big\}_{j=1}^{\infty}$ and $\{x_{k_j}\}_{j=1}^{\infty}$ such that $\big\{f_{k_j}\big\}_{j=1}^{\infty}$ converges uniformly to some function $f \in \big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$ and $\{x_{k_j}\}_{j=1}^{\infty}$ converges to some $a \in K$. We must also have $\{y_{k_j}\}_{j=1}^{\infty}$ converges to a since $d(x_{k_j}, y_{k_j}) < \dfrac{1}{k_j}$.

Since f is continuous at a,

$$\exists\, \delta > 0 \ni \|f(x) - f(a)\| < \frac{\varepsilon}{5} \qquad \text{if } x \in B(a,\delta) \cap K.$$

Moreover, since $\big\{f_{k_j}\big\}_{j=1}^{\infty}$ converges to f uniformly on K and $x_{k_j}, y_{k_j} \to a$ as $j \to \infty$, there exists $N > 0$ such that

$$\|f_{k_j}(x) - f(x)\| < \frac{\varepsilon}{5} \qquad \text{if } j \geqslant N \text{ and } x \in K$$

and

$$d(x_{k_j}, a) < \delta \quad \text{and} \quad d(y_{k_j}, a) < \delta \qquad \text{if } j \geqslant N.$$

As a consequence, for all $j \geqslant N$,

$$\begin{aligned} \varepsilon \leqslant \|f_{k_j}(x_{k_j}) - f_{k_j}(y_{k_j})\| &\leqslant \|f_{k_j}(x_{k_j}) - f(x_{k_j})\| + \|f(x_{k_j}) - f(a)\| \\ &\quad + \|f(y_{k_j}) - f(a)\| + \|f(y_{k_j}) - f_{k_j}(y_{k_j})\| < \frac{4\varepsilon}{5} \end{aligned}$$

which is a contradiction. □

Alternative proof of Lemma B.6. Suppose the contrary that B is not equi-continuous. Then there exists $\varepsilon > 0$ such that

$$\forall\, k \in \mathbb{N}, \exists\, x_k, y_k \in K \text{ and } f_k \in B \ni d(x_k, y_k) < \frac{1}{k} \text{ but } \|f_k(x_k) - f_k(y_k)\| \geqslant \varepsilon.$$

Since B is pre-compact in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$, there exists a subsequence $\{f_{k_j}\}_{j=1}^\infty$ converges to some function f in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$. By Proposition 7.46, $\{f_{k_j}\}_{j=1}^\infty$ converges uniformly to f on K; thus there exists $N_1 > 0$ such that

$$\|f_{k_j}(x) - f(x)\| < \frac{\varepsilon}{4} \quad \forall\, j \geqslant N_1 \text{ and } x \in K.$$

Since $f \in \mathscr{C}(K;\mathcal{V})$, by Theorem 4.49, f is uniformly continuous on K; thus

$$\exists\, \delta > 0 \ni \|f(x) - f(y)\| < \frac{\varepsilon}{4} \quad \text{if } d(x,y) < \delta \text{ and } x, y \in K.$$

Let $N = \max\big\{N_1, \big[\frac{1}{\delta}\big] + 1\big\}$, and $j \geqslant N$. Then $d(x_{k_j}, y_{k_j}) < \delta$ and this further implies that

$$\varepsilon \leqslant \|f_{k_j}(x_{k_j}) - f_{k_j}(y_{k_j})\| \leqslant \|f_{k_j}(x_{k_j}) - f(x_{k_j})\| + \|f(x_{k_j}) - f(y_{k_j})\| + \|f(y_{k_j}) - f_{k_j}(y_{k_j})\| < \frac{3\varepsilon}{4},$$

a contradiction. □

Corollary B.7. *Let (M, d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $K \subseteq M$ be a compact subset. If $\{f_k\}_{k=1}^\infty$ converges uniformly on K, then $\{f_k\}_{k=1}^\infty$ is equi-continuous.*

Example B.8. Corollary B.7 fails to hold if the compactness of K is removed. For example, let $\{f_k\}_{k=1}^\infty$ be a sequence of identical functions $f_k(x) = \dfrac{1}{x}$ on $(0,1)$. Then $\{f_k\}_{k=1}^\infty$ converges uniformly on $(0,1)$ but $\{f_k\}_{k=1}^\infty$ is not equi-continuous since none of f_k is uniformly continuous on $(0,1)$ which violates Remark B.2.

We have just shown that if $\{f_k\}_{k=1}^{\infty}$ converges uniformly on a compact set K, then $\{f_k\}_{k=1}^{\infty}$ must be equi-continuous. The inverse statement, on the other hand, cannot be true. For example, taking $\{f_k\}_{k=1}^{\infty}$ to be a sequence of constant functions $f_k(x) = k$. Then $\{f_k\}_{k=1}^{\infty}$ obviously does not converge, not even any subsequence. Therefore, we would like to study under what additional conditions, equi-continuity of a sequence of functions (defined on a compact set K) indeed converges uniformly. The following lemma is an answer to the question.

Lemma B.9. *Let (M, d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a Banach space, $K \subseteq M$ be a compact set, and $\{f_k\}_{k=1}^{\infty} \subseteq \mathscr{C}(K;\mathcal{V})$ be a equi-continuous sequence of functions. If $\{f_k\}_{k=1}^{\infty}$ converges pointwise on a dense subset E of K (that is, $E \subseteq K \subseteq \mathrm{cl}(E)$), then $\{f_k\}_{k=1}^{\infty}$ converges uniformly on K.*

Proof. Let $\varepsilon > 0$ be given. By the equi-continuity of $\{f_k\}_{k=1}^{\infty}$,

$$\exists\, \delta > 0 \ni \|f_k(x) - f_k(y)\| < \frac{\varepsilon}{3} \quad \text{if } d(x, y) < \delta,\ x, y \in K \text{ and } k \in \mathbb{N}.$$

Since K is compact, K is totally bounded; thus

$$\exists \{y_1, \cdots, y_m\} \subseteq K \ni K \subseteq \bigcup_{j=1}^{m} B\big(y_j, \frac{\delta}{2}\big).$$

By the denseness of E in K, for each $j = 1, \cdots, m$, there exists $z_j \in E$ such that $d(z_j, y_j) < \dfrac{\delta}{2}$. Moreover, $B\big(y_j, \frac{\delta}{2}\big) \subseteq B(z_j, \delta)$; thus $K \subseteq \bigcup\limits_{j=1}^{m} B(z_j, \delta)$. Since $\{f_k\}_{k=1}^{\infty}$ converges pointwise on E, $\{f_k(z_j)\}_{k=1}^{\infty}$ converges as $k \to \infty$ for all $j = 1, \cdots, m$. Therefore,

$$\exists\, N_j > 0 \ni \|f_k(z_j) - f_\ell(z_j)\| < \frac{\varepsilon}{3} \qquad \forall\, k, \ell \geqslant N_j\,.$$

Let $N = \max\{N_1, \cdots, N_m\}$, then

$$\|f_k(z_j) - f_\ell(z_j)\| < \frac{\varepsilon}{3} \qquad \forall\, k, \ell \geqslant N \text{ and } j = 1, \cdots, m\,.$$

Now we are in the position of concluding the lemma. If $x \in K$, there exists $z_j \in E$ such that $d(x, z_j) < \delta$; thus if we further assume that $k, \ell \geqslant N$,

$$\|f_k(x) - f_\ell(x)\| \leqslant \|f_k(x) - f_k(z_j)\| + \|f_k(z_j) - f_\ell(z_j)\| + \|f_\ell(z_j) - f_\ell(x)\| < \varepsilon\,.$$

By Proposition 7.7, $\{f_k\}_{k=1}^{\infty}$ converges uniformly on K. $\square$

Remark B.10. Corollary B.7 and Lemma B.9 imply that "a sequence $\{f_k\}_{k=1}^{\infty} \subseteq \mathscr{C}(K;\mathcal{V})$ converges uniformly on K if and only if $\{f_k\}_{k=1}^{\infty}$ is equi-continuous and pointwise convergent (on a dense subset of K)".

B.2 Compact sets in $\mathscr{C}(K;\mathcal{V})$

The next subject in this section is to obtain a (useful) criterion of determining the compactness (or pre-compactness) of a subset $B \subseteq \mathscr{C}(K;\mathcal{V})$ which guarantees the existence of a convergent subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ of a given sequence $\{f_k\}_{k=1}^{\infty} \subseteq B$ in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$.

Lemma B.11 (Cantor's Diagonal Process). *Let E be a countable set, $(\mathcal{V}, \|\cdot\|)$ be a Banach space, and $f_k : E \to \mathcal{V}$ be a sequence of functions. Suppose that for each $x \in E$, $\{f_k(x)\}_{k=1}^{\infty}$ is pre-compact in $\mathcal{V}$. Then there exists a subsequence of $\{f_k\}_{k=1}^{\infty}$ that converges pointwise on E.*

Proof. Since E is countable, $E = \{x_\ell\}_{\ell=1}^{\infty}$.

1. Since $\{f_k(x_1)\}_{k=1}^{\infty}$ is pre-compact in $(\mathcal{V}, \|\cdot\|)$, there exists a subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ such that $\{f_{k_j}(x_1)\}_{j=1}^{\infty}$ converges in $(\mathcal{V}, \|\cdot\|)$.

2. Since $\{f_k(x_2)\}_{k=1}^{\infty}$ is pre-compact in $(\mathcal{V}, \|\cdot\|)$, the sequence $\{f_{k_j}(x_2)\}_{j=1}^{\infty} \subseteq \{f_k(x_2)\}_{k=1}^{\infty}$ has a convergent subsequence $\{f_{k_{j_\ell}}(x_2)\}_{\ell=1}^{\infty}$.

Continuing this process, we obtain a sequence of sequences S_1, S_2, $\cdots$ such that

1. S_k consists of a subsequence of $\{f_k\}_{k=1}^{\infty}$ which converges at x_k, and

2. $S_k \supseteq S_{k+1}$ for all $k \in \mathbb{N}$.

Let g_k be the k-th element of S_k. Then the sequence $\{g_k\}_{k=1}^{\infty}$ is a subsequence of $\{f_k\}_{k=1}^{\infty}$ and $\{g_k\}_{k=1}^{\infty}$ converges at each point of E. □

The condition that "$\{f_k(x)\}_{k=1}^{\infty}$ is pre-compact in $\mathcal{V}$ for each $x \in E$" in Lemma B.11 motivates the following

Definition B.12. Let (M, d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $A \subseteq M$ be a subset. A subset $B \subseteq \mathscr{C}_b(A;\mathcal{V})$ is said to be ***pointwise*** $\begin{array}{c}\textbf{\textit{compact}}\\ \textbf{\textit{pre-compact}}\\ \textbf{\textit{bounded}}\end{array}$ if the set $B_x \equiv \big\{f(x) \,\big|\, f \in B\big\}$ is $\begin{array}{c}\text{compact}\\ \text{pre-compact}\\ \text{bounded}\end{array}$ in $(\mathcal{V}, \|\cdot\|)$ for all $x \in A$.

Example B.13. Let $f_k : [0,1] \to \mathbb{R}$ be given in Example B.5, and $B = \{f_k\}_{k=1}^{\infty}$. Then B is pointwise compact: for each $x \in [0,1]$, B_x is a finite set since if $f_k(0) = 0$ for all $k \in \mathbb{N}$, while if $x > 0$, $f_k(x) = 0$ for all k large enough which implies that $\#B_x < \infty$.

是時候可以來看 $\mathscr{C}(K;\mathcal{V})$ 裡面的 compact sets 有什麼等價條件了。首先我們先看何時 $B \subseteq \mathscr{C}(K;\mathcal{V})$ 是 compact set。給定一個函數列 $\{f_k\}_{k=1}^{\infty} \subseteq B$，我們想知道能不能找到一個在 sup-norm 下收斂的 subsequence $\{f_{k_j}\}_{j=1}^{\infty}$（即 sequentially compact）。由 Diagonal Process (Lemma B.11) 知，我們得在 K 中找一個稠密的子集合 E 使得 $\{f_k\}_{k=1}^{\infty}$ 在 E 上是 pointwise pre-compact（這個部份只保證了可以找到 subsequence 逐點收斂），然後加上 Lemma B.9 的幫助，馬上知道加上 equi-continuity 的條件之後，逐點收斂會變均勻收斂。因此，很自然地我們會要求 B 滿足 pointwise pre-compact 還有 equi-continuous 這兩個條件來證出 B 是 $\mathscr{C}(K;\mathcal{V})$ 中的 compact set。而在一個 compact set K 中能不能找到一個稠密子集合則是由下面這個 Lemma 所提供。

Lemma B.14. *A compact set K in a metric space (M,d) is separable; that is, there exists a countable subset E of K such that* $\text{cl}(E) = K$.

Proof. Since K is compact, K is totally bounded; thus $\forall\, n \in \mathbb{N}$, there exists $E_n \subseteq K$ such that

$$\#E_n < \infty \quad \text{and} \quad K \subseteq \bigcup_{y \in E_n} B\big(y, \frac{1}{n}\big).$$

Let $E = \bigcup\limits_{n=1}^{\infty} E_n$. Then E is countable by Theorem 0.20. We claim that $\text{cl}(E) = K$.

To see this, first by the definition of the closure of a set, $\text{cl}(E) \subseteq K$ (since K is closed). Let $x \in K$. Since $K \subseteq \bigcup\limits_{y \in E_n} B\big(y, \frac{1}{n}\big)$, $x \in B\big(y, \frac{1}{n}\big)$ for some $y \in E_n$. Therefore, $B\big(x, \frac{1}{n}\big) \cap E \neq \varnothing$ for all $n \in \mathbb{N}$. This implies that $x \in \bar{E} = \text{cl}(E)$. □

Theorem B.15. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a Banach space, $K \subseteq M$ be a compact set, and $B \subseteq \mathscr{C}(K;\mathcal{V})$ be equi-continuous and pointwise pre-compact. Then B is pre-compact in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$.*

Proof. We show that every sequence $\{f_k\}_{k=1}^{\infty}$ in B has a convergent subsequence. Since K is compact, there is a countable dense subset E of K (Lemma B.14), and the diagonal process (Lemma B.11) implies that there exists $\{f_{k_j}\}_{j=1}^{\infty}$ that converges pointwise on E. Since E is dense in K, by Lemma B.9 $\{f_{k_j}\}_{j=1}^{\infty}$ converges uniformly on K; thus $\{f_{k_j}\}_{j=1}^{\infty}$ converges in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$ by Proposition 7.46. □

Remark B.16. Lemma B.6 and Theorem B.15 imply that "a set $B \subseteq \mathscr{C}(K;\mathcal{V})$ is pre-compact if and only if B is equi-continuous and pointwise pre-compact". (That B is pre-compact implies that B is pointwise pre-compact is left as an exercise).

Corollary B.17. *Let (M,d) be a metric space, and $K \subseteq M$ be a compact set. Assume that $B \subseteq \mathscr{C}(K;\mathbb{R})$ is equi-continuous and pointwise bounded on K. Then every sequence in B has a uniformly convergent subsequence.*

Proof. By the Bolzano-Weierstrass theorem the boundedness of $\big\{f_k(x)\big\}_{k=1}^\infty$ implies that $\big\{f_k(x)\big\}_{k=1}^\infty$ is pre-compact for all $x \in E$. Therefore, we can apply Theorem B.15 under the setting $(\mathcal{V}, \|\cdot\|) = (\mathbb{R}, |\cdot|)$ to conclude the corollary. □

The following theorem provides how compact sets look like in $\mathscr{C}(K;\mathcal{V})$.

Theorem B.18 (The Arzelà–Ascoli Theorem)**.** *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a Banach space, $K \subseteq M$ be a compact set, and $B \subseteq \mathscr{C}(K;\mathcal{V})$. Then B is compact in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$ if and only if B is closed, equi-continuous, and pointwise compact.*

Proof. "$\Leftarrow$" This direction is conclude by Theorem B.15 and the fact that B is closed.

"$\Rightarrow$" By Lemma B.6 and the fact that compact sets are closed, it suffices to shows that B is pointwise compact. Let $x \in K$ and $\big\{f_k(x)\big\}_{k=1}^\infty$ be a sequence in B_x. Since B is compact, there exists a subsequence $\big\{f_{k_j}\big\}_{j=1}^\infty$ that converges uniformly to some function $f \in B$. In particular, $\big\{f_{k_j}(x)\big\}_{j=1}^\infty$ converges to $f(x) \in B_x$. In other words, we find a subsequence $\big\{f_{k_j}(x)\big\}_{j=1}^\infty$ of $\big\{f_k(x)\big\}_{k=1}^\infty$ that converges to a point in B_x. This implies that B_x is sequentially compact; thus B_x is compact. □

Example B.19. Let $f_k : [0,1] \to \mathbb{R}$ be a sequence of functions such that

(1) $|f_k(x)| \leqslant M_1$ for all $k \in \mathbb{N}$ and $x \in [0,1]$; (2) $|f_k'(x)| \leqslant M_2$ for all $k \in \mathbb{N}$ and $x \in [0,1]$.

Then $\{f_k\}_{k=1}^\infty$ is clearly pointwise bounded. Moreover, by the mean value theorem

$$\big|f_k(x) - f_k(y)\big| \leqslant M_2|x-y| \qquad \forall\, x, y \in [0,1], k \in \mathbb{N}$$

which implies that $\{f_k\}_{k=1}^\infty$ is equi-continuous. Therefore, by Corollary B.17 there exists a subsequence $\big\{f_{k_j}\big\}_{j=1}^\infty$ that converges uniformly on $[0,1]$.
Question: If assumption (1) of Example B.19 is omitted, can $\{f_k\}_{k=1}^\infty$ still have a convergent subsequence?

Answer: No! Take $f_k(x) = k$, then $\{f_k\}_{k=1}^{\infty}$ does not have a convergent subsequence (note that f_k is continuous and $f_k'(x) = 0$; that is, Assumption (2) is fulfilled).

Example B.20. We show that Assumption (1) of Example B.19 can be replaced by $f_k(0) = 0$ for all $k \in \mathbb{N}$.

Proof. (a) If $f_n(0) = 0$, then by the mean value theorem we have for all $x \in (0, 1]$ and $k \in \mathbb{N}$, $f_k(x) - f_k(0) = f_k'(c_k)(x - 0)$. Then Assumption (2) of Example B.19 implies that

$$\big|f_k(x) - f_k(0)\big| = \big|f_k'(c_k)\big|\big|x\big| \leqslant M_2|x| \leqslant M_2$$

which shows that $\{f_k\}_{k=1}^{\infty}$ is uniformly bounded by M_2.

(b) $\{f_k\}_{k=1}^{\infty}$ are equi-continuous (same proof as in Example B.19). □

B.3 Exercise

Problem B.1. Which of the following set B of continuous functions are equi-continuous in the metric space M? Are the closure $\bar{B}$ compact in M?

1. $B = \Big\{ \sin kx \,\Big|\, k = 0, 1, 2, \cdots \Big\}$, $M = \mathscr{C}([0, \pi]; \mathbb{R})$.
2. $B = \Big\{ \sin \sqrt{x + 4k^2\pi^2} \,\Big|\, k = 0, 1, 2, \cdots \Big\}$, $M = \mathscr{C}_b([0, \infty); \mathbb{R})$.
3. $B = \Big\{ \dfrac{x^2}{x^2 + (1 - kx)^2} \,\Big|\, k = 0, 1, 2, \cdots \Big\}$, $M = \mathscr{C}([0, 1]; \mathbb{R})$.
4. $B = \Big\{ (k + 1)x^k(1 - x) \,\Big|\, k \in \mathbb{N} \Big\}$, $M = \mathscr{C}([0, 1]; \mathbb{R})$.

Problem B.2. Let (M, d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed space, and $A \subseteq M$ be a subset. Suppose that $B \subseteq \mathscr{C}_b(A; \mathcal{V})$ be equi-continuous. Prove or disprove that $\mathrm{cl}(B)$ is equi-continuous.

Problem B.3. Let $f_k : [a, b] \to \mathbb{R}$ be a sequence of differentiable functions such that $f_k(a)$ is bounded and $|f_k'(x)| \leqslant M$ for all $x \in [a, b]$ and $k \in \mathbb{N}$. Show that $\{f_k\}_{k=1}^{\infty}$ contains an uniformly convergent subsequence. Must the limit function differentiable?

Problem B.4. Let $\mathscr{C}^{0,\alpha}([0, 1]; \mathbb{R})$ denote the "space"

$$\mathscr{C}^{0,\alpha}([0, 1]; \mathbb{R}) \equiv \Big\{ f \in \mathscr{C}([0, 1]; \mathbb{R}) \,\Big|\, \sup_{x,y \in [0,1]} \frac{|f(x) - f(y)|}{|x - y|^{\alpha}} < \infty \Big\},$$

where $\alpha \in (0,1]$. For each $f \in \mathscr{C}^{0,\alpha}([0,1];\mathbb{R})$, define

$$\|f\|_{\mathscr{C}^{0,\alpha}} = \sup_{x\in[0,1]} |f(x)| + \sup_{\substack{x,y\in[0,1]\\ x\neq y}} \frac{|f(x)-f(y)|}{|x-y|^\alpha}.$$

1. Show that $\big(\mathscr{C}^{0,\alpha}([0,1];\mathbb{R}), \|\cdot\|_{\mathscr{C}^{0,\alpha}}\big)$ is a complete normed space.

2. Show that the set $B = \big\{f \in \mathscr{C}([0,1];\mathbb{R}) \,\big|\, \|f\|_{\mathscr{C}^{0,\alpha}} < 1\big\}$ is equi-continuous.

3. Show that $\mathrm{cl}(B)$ is compact in $\big(\mathscr{C}([0,1];\mathbb{R}), \|\cdot\|_\infty\big)$.

Problem B.5. Given $f:\mathbb{R}\to\mathbb{R}$ a continuous periodic function of period 1; that is, $f(x+1) = f(x)$ for all $x \in \mathbb{R}$, and $x_1,\cdots,x_m \in [0,1]$ arbitrary m points, define a new function

$$I(f;x_1,\cdots,x_m)(x) = \frac{1}{m}\big[f(x+x_1)+\cdots+f(x+x_m)\big] \qquad \forall\, x\in\mathbb{R}.$$

Prove that the set

$$B = \big\{I(f;x_1,\cdots,x_m)\,\big|\, x_1,\cdots,x_m \in [0,1], m\in\mathbb{N}\big\}$$

is uniformly bounded and equi-continuous in the space $\mathscr{C}([0,1];\mathbb{R})$. Moreover, show that the derived set $B' = \Big\{\int_0^1 f(x)dx\Big\}$; that is, the derived set of B consists of one single function which is a constant function $y = \int_0^1 f(x)dx$.

Problem B.6. Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a Banach space, $K \subseteq M$ be compact, and $\{f_k\}_{k=1}^\infty \subseteq \mathscr{C}(K;\mathcal{V})$ be a sequence of continuous functions. Suppose that for all $x \in K$, if $\{x_k\}_{k=1}^\infty, \{y_k\}_{k=1}^\infty \subseteq K$ and $\lim\limits_{k\to\infty} x_k = \lim\limits_{k\to\infty} y_k = x$, the limits $\lim\limits_{k\to\infty} f_k(x_k)$ and $\lim\limits_{k\to\infty} f_k(y_k)$ exist and are identical. Show that $\{f_k\}_{k=1}^\infty$ converges uniformly on K. How about if K is not compact?

Problem B.7. Assume that $\{f_k\}_{k=1}^\infty$ is a sequence of monotone increasing functions on $\mathbb{R}$ with $0 \leqslant f_k(x) \leqslant 1$ for all $x\in\mathbb{R}$ and $k\in\mathbb{N}$.

1. Show that there is a subsequence $\{f_{k_j}\}_{j=1}^\infty$ which converges pointwise to a function f (This is usually called the ***Helly selection theorem***).

2. If the limit f is continuous, show that $\{f_{k_j}\}_{j=1}^\infty$ converges uniformly to f on any compact set of $\mathbb{R}$.

Bibliography

[1] George Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, 2(4):303–314, 1989. Universal approximation for sigmoidal activation functions. 7.73

[2] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989. Early universal approximation results for single hidden layer networks. 7.73

[3] Moshe Leshno, Vladimir Y. Lin, Allan Pinkus, and Shimon Schocken. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Networks*, 6(6):861–867, 1993. Characterization of activation functions for universal approximation. 7.73

[4] Allan Pinkus. Approximation theory of the mlp model in neural networks. *Acta Numerica*, 8:143–195, 1999. Comprehensive survey of approximation theory for neural networks. 7.73

Index