從有限到無限-Delta 函數面面觀

陳卓老師 文

§1. 理論物理一代宗師狄拉克氏(P. A. M. Dirac 1902-1984)在建構量子力學之普遍理論時引入一種奇異函數(singular function)，稱之為 δ-function(δ 及希臘之中的 Delta，相當於英文字母 D)。狄氏對 δ-函數並沒有嚴格的定義，但描述了這個奇異函數的基本性質，即

$$\int_{-\infty}^{\infty}\delta(x)dx=1 \quad （保元性）(\mathrm{I})$$

且$\delta(\mathrm{x})=0$，對 $\mathrm{x}\neq 0$

在$x=0$的極窄臨近，即$|x|<\epsilon$，而ϵ是一個極小微量，$\delta(\mathrm{x})$是一個奇峰凸起的函數，使它在這微量區間內的積分(面積)等於 1，至於$\delta(\mathrm{x})$在這區域內的具體形狀並不重要[1]

根據這個想法，$\delta(\mathrm{x})$會有下列性質：

若$f(x)$是一個正常函數，在$x=0$臨近視連續且可為分者，則

(i)

$$\int_{-\infty}^{\infty}f(x)\delta(x)\,dx=f(0)$$

[1]參考：狄氏所著：Principles of Quantum Mechanics 4thEd. §15, p58

(ii)若將 δ (x)稍微推廣，考慮一任一點x_o，則

$$\int_{-\infty}^{\infty}\delta(\mathrm{x}-\mathrm{x}_0)\,dx=1$$

且 $\delta(\mathrm{x}-\mathrm{x}_0)=0$當$x\neq x_o$，並有

$$\int_{-\infty}^{\infty}f(x)\delta(x-x_0)dx=f(x_0)$$

此處 f(x)這函數再 x_0 這點的鄰近須為連續及可微分者。

(iii)再推廣，若 a>0，則

$$\delta[a(x-x_0)]=\frac{1}{a}\delta(x-x_0)$$

這個關係是的來源是微積分中的變數替代法。即令 $u=ax$，則 $u_0=ax_0$，$du=adx$，$dx=\frac{1}{a}du$

(iv) 再推廣，若 $f(x)$ 在 x_0 為一連續且可微分函數，且 $f'(x_0)\neq 0$ ，則

$$\delta[f(x)]=\left|\frac{df(x)}{dx}\right|^{-1}\delta(x-x_0)$$

(v) $\delta'(x)$ 之定義

$$\int_{-\infty}^{\infty}\delta'(x)f(x)dx=-f'(0)$$

說明： $\delta(x)$ 雖然是一個 Singular fundtion，但是可以視為某個行為良好函數數列的極限，這些函數都可以微分，所以 $\delta'(x)$ 亦有意義。由積分中的部分

積分

$$\int_a^b \frac{d}{dx}[f(x)g(x)]dx = [f(x)g(x)]_a^b$$

$$= \int_a^b f'(x)g(x)dx + \int_a^b g'(x)f(x)dx$$

在上式中，若取 $g(x) = \delta(x)$，而 $[a, b] \to [-\infty, \infty]$。則除 $x = 0$ 的鄰近外 $g(x) = 0$。故

$$[f(x)g(x)]_a^b = 0$$

$$\int_a^b g'(x)f(x)dx = -\int_a^b f'(x)g(x)dx$$

$$\int_a^b \delta'(x)f(x)dx = -\int_a^b f'(x)\delta(x)dx$$

只要 $[a, b]$ 包含 $x = 0$ 這一點，則故

$$\int_a^b \delta'(x)f(x)dx = -f'(0)$$

我們應注意，積分的上下限 $[a, b]$ 只要包含 $x = 0$ 這一點，對 $\delta(x)$ 而言，便與 $[-\infty, \infty]$ 沒什麼差別。

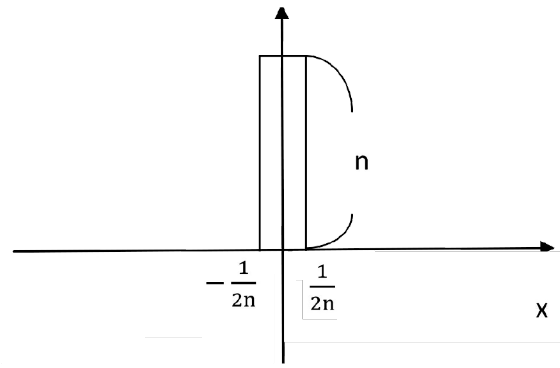
§2. δ函數之數列表示法

由於δ函數在量子力學應用中的重要性，又在物理及工程中的應用範圍逐漸擴大，受到「應用數學」家們的普遍重視，遂成為應用解析學中的一個課題。許多程度高的應用數學課本中都有專門章節討論δ函數[2]。

在此處，我們可以把δ函數視為某個行為良好的函數的數列 (Sequense of Well-behaved Functions) 之極限；所謂「行為良好者」，通常就是指可以微分二次以上的函數（連續）。以下我們先列出幾個大家都熟悉的δ函數的「數列表示」。

(1)

$$\delta_n(x) = \begin{cases} 0, & x < -\frac{1}{2n} \\ n, & -\frac{1}{2n} < x < \frac{1}{2n} \\ 0, & x > \frac{1}{2n} \end{cases}$$



[2] (1) Arfken and Weber 所著：Mathematical Methods for Physicists, 5ED, 第一章, 15 節, p.84；或 (2) P. Dennery and A. Krzywicki 所著 Mathematics for Physicists, 第三章, 13 節, p.225

這個數列表示法式最原始的表示法。當 $n \gg 1$ 時 $\frac{1}{n} \equiv \epsilon_n$ 即越來越小，最後趨近於零，而 $n = \frac{1}{\epsilon}$ 則趨向無窮大，而 $\int_{-\infty}^{\infty} \delta_n(x)dx = 1$，無論 n 有多大。

這個形式的表示法在「電路分析」(Network Analysis) 中被稱為 unit-pulse function （即一個脈衝訊號，例如電流 $i_n(t)$ ，其積分 $\int_{-\infty}^{\infty} i_n(t)dt = 1$，即通過的總電量 $\Delta Q = \int_{-\infty}^{\infty} i_n(t)dt = 1$）

這種函數的缺點是它的不連續性，在 $x = \pm\frac{1}{2n}$ 這兩點不連續。但若有一連續函數 $f(x)$ 它在 $x = 0$ 鄰近為連續，則

$$\int_{-L}^{L} f(x)\delta_n(x)dx = \int_{-\frac{1}{2n}}^{\frac{1}{2n}} f(x) \cdot n\,dx$$

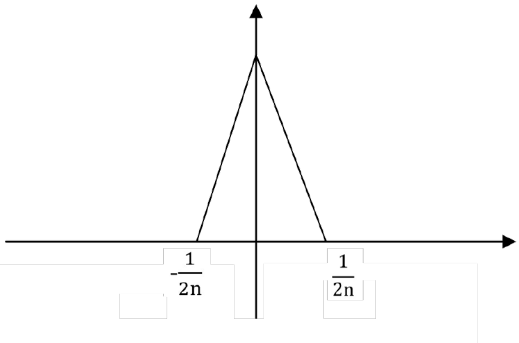
$$= n \cdot \int_{-\frac{1}{2n}}^{\frac{1}{2n}} f(x)dx$$

$$= n \cdot f(0) \cdot \left[\frac{1}{2n} - \left(-\frac{1}{2n}\right)\right] = f(0)$$

無論如何，這種形態的「函數」不能被視為「行為良好」的函數。

(2) 若將 (1) 式稍微變形

$$\delta_n(x) = \begin{cases} 0, \ x \le -\frac{1}{2n} \\ 4n^2x + 2n, \ \ -\frac{1}{2n} \le x < 0 \\ -4n^2x + 2n, \ 0 < x \le \frac{1}{2n} \\ 0, \ x > \frac{1}{2n} \end{cases}$$



這個 $\delta_n(x)$ 的數列表示的形狀即如圖中所示，與前面的矩形函數稍微不同，它是一個連續函數。在 $x = \pm\frac{1}{2n}$ 及 $x = 0$ 這三點都連續但不可微分。它還是一個 unit-pulse function，滿足 δ-function 的基本要求，而且更像一個「奇峰凸起」的形狀。但它們尚不能被視為「行為良好」的函數數列。

(3) 高斯型 (Gaussian) 函數數列

$$\delta_n(x) = \frac{n}{\sqrt{\pi}} e^{-n^2x^2}$$

這個形式的函數數列在機率論中被稱為高斯分佈 (Gaussian Distribution)，或常態分佈 (Normal Distribution)。很明

顯這是一種「行為良好」的函數數列，所以最受應用者的歡迎，在許多場合都被抬出來作為論證時的「工具」(利器)。這些函數 $\delta_n(x)$ 都滿足

$$\int_{-\infty}^{\infty} \delta_n(x)\, dx = 1$$

而且 e^{-x^2/a^2} 這種函數隨 x 之增加，函數值下降非常快(以a為單位，$x = na$，$e^{-x^2/a^2} = e^{-n^2} = 1/e^{n^2}$)，在 $x = 0$ 這一點 $\frac{n}{\pi}e^{-n^2x^2} = \frac{n}{\pi} \to \infty\ as\ n \to \infty$。

Gaussian Distribution 的半寬度 (Half-Width) $\varDelta$，可定義為使

$$\delta_n(x) = \frac{n}{\pi}e^{-n^2x^2} = \frac{1}{2}\left(\frac{n}{\pi}\right)$$

的 $\varDelta$ 值，故

$$e^{-n^2\varDelta^2} = \frac{1}{2}$$

即

$$n^2\varDelta^2 = \ln 2$$

或

$$\varDelta = \frac{\ln 2}{n}$$

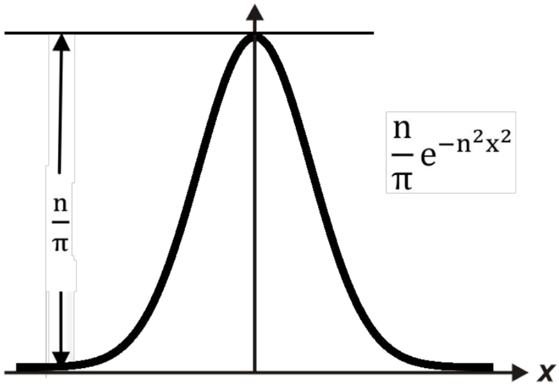
若取 $n = 10^4$， $\varDelta = 0.69\times10^{-4}$

故

$$\varDelta_n \propto \frac{const.}{n}$$

總而言之，隨著 n 增加，這分佈函數的「寬度」迅速下降，其關係式為

$$\varDelta = \frac{1}{n}\ln 2$$



(4) 勞倫茲型 (Lorenzian)

$$\delta_n(x) = \frac{n}{\pi}\cdot\frac{1}{1+n^2x^2}\text{，}\quad n\text{大正整數}$$

這一型分佈函數隨 n 增加而下降的趨勢比較和緩，它們亦滿足歸元關係式

$$\int_{-\infty}^{\infty} \delta_n(x)dx = 1,\quad for\ all\ n$$

在 $x = 0$ 處 $\delta_n(0) = \frac{n}{\pi}$，故其「高度」仍與 n 成正比。其半寬度可表示為

$$\frac{n}{\pi}\cdot\frac{1}{1+n^2\varDelta^2} = \frac{1}{2}\left(\frac{n}{\pi}\right)$$

即

$$\frac{1}{1+n^2\varDelta^2} = \frac{1}{2}$$

或

$$1 + n^2\varDelta^2 = 2$$

即

$$\Delta = \frac{1}{n}$$

雖然 Lorenzian Distribution 的半寬度 Δ 與 Gaussian Distribution 相當，但因為它的尾巴拖得很長，作為 δ 函數的數列表示函數並不式很理想。（不如 Gaussian Distribution）

(5) 三角函數型 （又稱 Fraunhoffer 型）

$$\delta_n(x) = \frac{\sin nx}{\pi x} = 2\pi \int_{-n}^{n} e^{-ikx} dk$$

這一型函數源自 Fraunhoffer Diffraction 中光強度計算，是波動理論中會「自然發生」的函數，有其特別重要性。這一型函數的「歸元性」來自積分形式

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx = \pi$$

◆ 以上所列 Gaussian 和 Lorentian 數列表示法都和 Fourier Transorm 相通，故通常把$\delta(t)$寫成

$$\delta(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega t} d\omega$$

或

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx} dk$$

為了說明其間關係，我們可以引用下面的積分等式

$$(A) \qquad \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t - \frac{\epsilon}{4}\omega^2} d\omega = \frac{1}{\sqrt{\pi\epsilon}} e^{-\frac{t^2}{\epsilon}}$$

若令 $\frac{1}{\epsilon} = n^2$ ，上式左側或為 Gaussian Rep.

$$(B) \qquad \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t - \epsilon|\omega|} d\omega = \frac{1}{\pi}\frac{\epsilon}{t^2 + \epsilon^2}$$

若令$\epsilon = \frac{1}{n}$ ，上式左側即成 Lorentian Distribution。

§3. 函數論與向量空間

在數學的應用中，尤其是牽涉到數值計算的問題，人們常需要把某一個已知（或未知）函數用一組大家都熟悉的函數來展開。最常見的便是泰勒展開式（包括 Maclaurin 展開），其他較常見的有 Fourier 展開（三角函數展開），勒襄德 Polynomial (Legendre)，厄米德多項式 (Hermite Polynomial)等等。這些展開需要一個共同的理論基礎，而這理論基礎便是：任何連續函數（包括 piece-wise discontinous function with finite discontinuity，即不連續處之不連續幅度為有限者，如下圖所式）可以用多項式來逼近。

若連續函數之定義域為有限區間 [a,b] 則很容易做變數變換，定義

$$\bar{x} = \frac{1}{b-a}(x-a)$$

則 $\bar{x}$ 的定義域就變為 [0,1]

任何有關連續函數 $f(x)$，$x \in [0,1]$ 的論證若成立，則經變數變換後，對於定義在 $\bar{x} \in [a,b]$上的函數亦成立。因此我們只要把理論探討的注意力集中在定義於[0,1]這個封閉區間上即可。

◆ 定義在 [0,1] 區間上的連續函數之向量表示法

考慮一「連續」函數 $f(x)$，若將 [0,1] 這區域分割成 N 等分（N 很大），若定義

$$x_k = \frac{k}{N} \ ,k = 0,1,2,\cdots,N$$

當 N 很大時，$x_{k+1} - x_k = \frac{1}{N}$ 便很小。對於一個「連續」函數而言，$f(x_{k+1})$ 與 $f(x_k)$ 之差別也就很小，而且在 x_k與 x_{k+1}之間 $f(x)$的函數值亦不會與 $f(x_k)$ 及 $f(x_{k+1})$ 差很多。要不然$f(x)$就不是連續函數了。例如：若要在 (x_k, x_{k+1})之間 $f(x)$ 的函數值與 $f(x_k)$ 之差別不超過 $\epsilon = 10^{-4}$，我們一定可以找到一個 N，也許是10^6，即 $|x_{k+1} - x_k| = 10^{-6}$，使$|f(x) - f(x_k)| < \epsilon = 10^{-4}$

因此，只要 N 夠大，一組函數值 $\{f(x_k),\ k = 0,1,2,3,\cdots,N\}$ 便足以「表現」f 這個函數。套句數學術語，這$(N+1)$個函數值便足以將 f 這個函數「支撐」起來。所以我們便可將 f 表示為一個向量

$$|f\rangle \equiv \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_N \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \\ \vdots \\ f(x_N) \end{bmatrix}$$

而構成一個 N+1 維的向量空間。($|f\rangle$ 這個符號是 Dirac 提倡的）任何連續函數都可以用這種 $(N+1)$ 維向量空間中的向量來表示。其加法可以表示為：

$$h(x) = \alpha f(x) + \beta g(x) \Rightarrow h = \alpha f + \beta g$$

$$|h\rangle = \begin{bmatrix} \alpha f_0 + \beta g_0 \\ \alpha f_1 + \beta g_1 \\ \vdots \\ \alpha f_N + \beta g_N \end{bmatrix} = \alpha|f\rangle + \beta|g\rangle$$

而這兩個函數 $|f\rangle$ 與 $|g\rangle$ 之內積 (Inner Product)之自然定義為

$$\langle f|g\rangle = \frac{1}{N+1}\sum_{k=0}^{N} f_k^* g_k$$ [3]

很明顯，在取 $N \to \infty$ 時

[3] $\frac{1}{N+1}$ 這個因子可以保證，在N很大時，$\langle f|g\rangle$ 之數值與 N 無關；又若 f 及 g 皆為實數，則 $*$ 這個符號便不需要

$$\langle f|g\rangle = \int_0^1 f^*(x)g(x)dx$$

◆ 在上述向量空間中，很自然的一組基底便是

$$|\rho_k\rangle = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1\text{ , } (the\ k_{th}\ component) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$, k = 0,1,2,\cdots,N$

但這樣定義出來的基底組 $\{|\rho_k\rangle,\ k = 0,1,2,\cdots,N\}$ 並不滿足「歸一化」(normalization)的關係，而成為

$$\langle\overrightarrow{\rho_k}|\overrightarrow{\rho_k}\rangle = \frac{1}{N+1} \to 0 \quad as\ \ n \to \infty$$

我們可以稍微修正，而定義

$$|x_k\rangle = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ N+1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

然而這樣定義出來的基底組一不滿足歸一化（保元）條件，但有以下的優點

(i) $\langle x_j|x_k\rangle = (N+1)\delta_{jk}$

$$= (N+1)\begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases}$$

(ii) 如果 f 是一個正常函數（連續而可微分），則

$$\langle x_k|f\rangle = f(x_k)$$

(iii) $$\frac{1}{N+1}\sum_a |x_k\rangle\langle x_k| = \mathbb{1}$$

上式右側之 $\mathbb{1}$ 即為 Identity Operator。這是一個很重要的關係式，可以由下列事實彰顯。

(a) $$|f\rangle = \frac{1}{N+1}\sum_{k=0}^{N} |x_k\rangle\langle x_k|f\rangle$$

$$= \frac{1}{N+1}\sum_{k=0}^{N} f(x_k)|x_k\rangle = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ \vdots \\ f(x_k) \end{bmatrix}$$

(b) $$\langle f|g\rangle = \frac{1}{N+1}\sum_{k=0}^{N} \langle f|x_k\rangle\langle x_k|g\rangle$$

$$\frac{1}{N+1}\sum_{k=0}^{N} f^*(x_k)g(x_k)$$

由以上這些關係式可以看出，$\langle x_i|x_j\rangle$ 事實上就是δ-function 的根源，即

$$\langle x|x'\rangle = \delta(x-x') = \begin{cases} 0, & x \neq x' \\ \infty, & x = x' \end{cases}$$

而

$$\langle x|f\rangle = \frac{1}{N+1}\sum_k \langle x|x_k\rangle\langle x_k|f\rangle$$

取 $N \to \infty$之極限後即成

$$f(x) = \int \delta(x-x')f(x')dx'$$

$$here\begin{cases} x' \to x_k \\ \dfrac{1}{N+1} \to dx' \end{cases}$$

以上這些論證雖然堪稱清楚明白但仍略有粗糙之嫌，然而可以將論證再精緻化，其辦法是運用俄國數學家 Berstein 氏為了重新證明 Stone-Weierstrass Theorem 而建構的多項式，後人稱之為 Berstein Polynomials。[4]

§4. 二項式分布與有限區間上的 δ-function

首先考慮二項式定理

$$(a+b)^N = \sum_{k=0}^{N} C_k^N a^k b^{N-k}$$

$$C_k^N = \frac{N!}{k!\,(N-k)!}$$

若取 $a = x$, $b = (1-x)$, 則 $a + b = 1$, 於是

$$1 = \sum_{k=0}^{\infty} C_k^N x^k (1-x)^{N-k}\ , \qquad x \in [0,1]$$

$$= \frac{1}{N+1}\sum_{k=0}^{\infty} (N+1) C_k^N x^k (1-x)^{N-k}$$

上式中左側 $\sum_k$ 之中每一項皆為一多項式，其中第 k 項可寫成

$$B_N^k(x) = (N+1) C_k^N x^k (1-x)^{N-k}$$
$$, k = 0,1,2,\cdots,N$$

不妨稱之為 Binomial Polynomial。他們皆滿足（k=0 與 k=N 除外）

(i) $B_N^k(0) = B_N^k(1) = 0$, $k = 1,2,\cdots,(N-1)$

(ii) $\int_0^1 B_N^k(x)dx = 1$ （歸一化）

(iii) $B_N^k(x)$ 在 $x = x_k = \frac{k}{N}$ 處有一尖峰，其高度約為 $\sqrt{N}$ 而寬度約

以上 (ii) 或可由積分直接驗證，第 (iii) 式則可將 $B_N^k(x)$ 微分，求得其極大值及拐點之位置，再求出在這些點的函數值，並分析函數之形狀，即可知其為真。

◆ 演算細節

(ii)

$$\int_0^1 B_N^k(x)dx = \frac{(N+1)!}{k!\,(N-k)!}\int_0^1 x^k (1-x)^{N-k} dx$$

這個積分可表示為 Beta-Function

[4] George E. Simmons, Introduction to Topology ans Modern Analysis, Ch7, p.153

$$B(a,b) \equiv \int_0^1 x^{a+1}(1-x)^{b+1}dx$$

$$= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

取$a = k+1, b = N-k+1$，則

$$\int_0^1 x^k(1-x)^{N-k}dx$$

$$= \frac{\Gamma(k+1)\Gamma(N-k+1)}{\Gamma(N-k+k+2)}$$

$$= \frac{\Gamma(k+1\,)\Gamma(N-k+1)}{\Gamma(N+2)} = \frac{k!\,(N-k)!}{(N+1)!}$$

故 $\int_0^1 B_N^k(x)dx = 1 \qquad for\ all\ k$

(iii) 再計算 $B_N^k(x)$ 之微分

$$\frac{d}{dx}B_N^k(x) =$$

$$\frac{(N+1)!}{k!\,(N-k)!}x^{k-1}(1-x)^{N-k-1}(k-Nx)$$

故它的極大值位於 $x = \frac{k}{N} = x_k$

再微分一次，求出

$$\frac{d^2}{dx^2}B_N^k(x) =$$

$$\frac{(N+1)!}{k!\,(N-k)!}x^{k-2}(1-x)^{N-k-2}$$

$$\cdot [N(N-1)x^2 - 2k(N-1)x + k(k-1)]$$

故 $\frac{d^2}{dx^2}B_N^k(0) = 0$ 的位置在

$$N(N-1)x^2 - 2k(N-1)x + k(k-1) = 0$$

即 $x = \frac{k}{N} \pm \frac{1}{\sqrt{N-1}}\sqrt{\frac{k}{N}(1-\frac{k}{N})}$

所以兩個拐點之間距為

$$\frac{2}{\sqrt{N-1}}\sqrt{\frac{k}{N}(1-\frac{k}{N})} \propto \frac{2}{\sqrt{N}}$$

（若 $N \to \infty$，而 $\frac{k}{N}$ 為有限值）

$B_N^k(x)$ 在 $x = x_k$ ，即它的頂峰時，其數值可以用 Stirling's Formula，即

$$N! \approx \sqrt{2\pi}N^{N+\frac{1}{2}}e^{-N}$$

做近似計算，其結果為

$$B_N^k(x) \approx \frac{\sqrt{N}}{2\pi}\left[\frac{k}{N}\left(1-\frac{k}{N}\right)\right]^{-\frac{1}{2}}$$

$$\propto \sqrt{N} \qquad (\frac{k}{N}\text{為有限值})$$

從以上分析可知，$B_N^k(x)$這個多項式其實就是在 [0,1]這區間上的 δ-函數（當 $N \gg 1$），所以我們不妨將它改寫（定義）為$D_N(x, x_k)$其中x_k就是那個尖峰位置 $x_k = \frac{k}{N}$ ，$k = 1,2,,\cdots,(N-1)$。而且可以期待，對任何定義在 [0,1] 上的正常函數 $f(x)$

$$f(x) \cong \frac{1}{N+1}\sum_{k=0}^{N} D_N(x, x_k)f(x_k) \quad --(I)$$

其中

$$D_N(x,x_k)=\frac{(N+1)!}{k!\,(N-k)!}(1-x)^{N-k}x^k$$

在取連續極限 (Continuous Limit, 即 $N\to\infty$)後，把x_k用x'替代，即可得

$$f(x)=\int_0^1 D(x,x')f(x')dx'$$

其中$D(x,x')$即為δ-function $\delta(x-x')$

在 N 仍為一有限數時，上面(I)式中等號右側仍為一個多項式，它與函數$f(x)$有關，有時被稱為"與$f(x)$連結的 Berstein Polynomial" (Berstein Polynomial associated with $f(x)$)。

以下我們將給出一個「嚴格的證明」(Rigorous Prooof)

$$\lim_{n\to\infty}\frac{1}{N+1}\sum_{k=0}^{N}\frac{(N+1)!}{k!\,(N-k)!}x^k(1-x)^{N-k}f\left(\frac{k}{N}\right)=f(x)$$

證明：要證明以上等式成立，就是要證明：

對於任意設定的微量$\epsilon>0$，可以找出一個正整數N_0，使所有大於N_0的正整數 N 都滿足

$$|f_N(x)-f(x)|<\epsilon$$

而

$$f_N(x)\equiv\sum_{k=0}^{N}\frac{(N+1)!}{k!\,(N-k)!}x^k(1-x)^{N-k}f(\frac{k}{N})$$

這個證明需要利用$f(x)$在$[0,1]$這區間上的連續性。即若x及x'兩點都在$[0,1]$區間之內，則對於任意微正量$\epsilon(>0)$，必可找出一個微小區間，即有$\delta>0$

在$|x-x'|<\delta$之內，$f(x)$與$f(x')$之差必小於$\frac{\epsilon}{2}$，即$|f(x)-f(x')|<\frac{\epsilon}{2}$

意即，只要$|x-x'|$夠窄，則這兩點函數值的差別(之絕對值)就一定很小。

現在考慮特定的$x\in[0,1]$，讓我們把$[0,1]$這個區間分成兩部份，實際上是把所有的$\{x_k\}$分成兩部分：(i)滿足$|x-x_k|<\delta$者，(ii)不滿足這關係式者。

凡是屬於第(i)類的x_k，皆滿足$|f(x)-f(x_k)|<\epsilon$，而屬於第(ii)類者則未必。

又因為

$$\sum_{k=0}^{N}C_k^N\,x^k(1-x)^{N-k}=1$$

故

$$\sum_{k=0}^{N}C_k^N\,x^k(1-x)^{N-k}f(x)=f(x)$$

故

$$\left|\sum_{k=0}^{\infty}C_k^N\,x^k(1-x)^{N-k}f(x_k)-f(x)\right|$$

$$= \left| \sum_{k=0}^{\infty} C_k^N x^k (1-x)^{N-k} \big(f(x_k) - f(x)\big) \right|$$

$$\leq \sum_{k=0}^{\infty} C_k^N x^k (1-x)^{N-k} |f(x) - f(x_k)|$$

$$= \sum\nolimits_k' C_k^N x^k (1-x)^{N-k} |f(x) - f(x_k)|$$

$$+ \sum\nolimits_k'' C_k^N x^k (1 - x)^{N-k} |f(x) - f(x_k)|$$

以上$\sum_k'$ 表示對第一類x_k各項求和，$\sum_k''$表示對第一類x_k各項求和

對第一類諸x_k，$|f(x) - f(x_k)| < \frac{\epsilon}{2}$

故

$$\sum\nolimits_k' C_k^N x^k (1-x)^{N-k} |f(x) - f(x_k)|$$

$$\leq \frac{\epsilon}{2} \sum\nolimits_k' C_k^N x^k (1-x)^{N-k}$$

$$\leq \sum_{k=0}^{N} C_k^N x^k (1-x)^{N-k} = \frac{\epsilon}{2}$$

剩下的工作是要證明

$$\sum\nolimits_k'' C_k^N x^k (1-x)^{N-k} |f(x) - f(x_k)| < \frac{\epsilon}{2}$$

我們應注意，在此類x_k中，$|x - x_k|^2 \geq \delta^2$

故$\frac{|x-x_k|^2}{\delta^2} \geq 1$，但因$f(x)$在$[0,1]$區間內為連續函數，$f(x)$的絕對值必有上限，設$K$是一個$|f(x)|$的上限，則$|f(x) - f(x_k)| < 2K$

故

$$\sum\nolimits_k'' C_k^N x^k (1-x)^{N-k} |f(x) - f(x_k)|$$

$$\leq 2K \sum\nolimits_k'' C_k^N x^k (1-x)^{N-k}$$

$$\leq 2K \sum\nolimits_k'' C_k^N x^k (1-x)^{N-k} \frac{|x - x_k|^2}{\delta^2}$$

$$\leq \frac{2K}{\delta^2} \sum\nolimits_k'' C_k^N x^k (1-x)^{N-k} |x - x_k|^2$$

$$\leq \frac{2K}{\delta^2} \sum_{k=0}^{N} C_k^N x^k (1-x)^{N-k} (x - \frac{k}{N})^2$$

(注意 sum over all k > sum over partial k’s)

即

$$\sum\nolimits_k'' C_k^N x^k (1-x)^{N-k} |f(x) - f(x_k)|$$

$$\leq \frac{2K}{\delta^2} \sum_{k=0}^{N} C_k^N x^k (1-x)^{N-k} (x - \frac{k}{N})^2$$

但以上右側的 summation 事實上是 Binominal Distribution 的 Mean-Square Deviation，是一個可計算的量，它等於$\frac{x(1-x)}{N}$(證法細節見後)，$\frac{x(1-x)}{N}$的極大值發生在$x = 1/2$，

$$\frac{\frac{1}{2}(1 - \frac{1}{2})}{N} = \frac{1}{4N}$$

故

$$\frac{2K}{\delta^2}\sum_{k=0}^{N} C_k^N x^k (1-x)^{N-k}(x-\frac{k}{N})^2 \le \frac{1}{4N}$$

把這個關係式代回關於$\sum_k''$ 的關係式，即得

$$\sum\nolimits_k'' C_k^N x^k (1-x)^{N-k}|f(x)-f(x_k)|$$

$$\le \frac{2K}{\delta^2}\frac{1}{4N} = \frac{K}{2N}(\frac{1}{\delta^2})$$

對於已經設定的$\epsilon(>0)$，δ亦可以求出

如果我們取一個夠大的N_0，使它滿足

$$\frac{K}{2N_0}(\frac{1}{\delta^2}) < \frac{\epsilon}{2} \quad (\text{或}N_0 > \frac{K}{\delta^2\epsilon})$$

則$\sum_k''$ 這一項變小於$\frac{\epsilon}{2}$，於是

$$\sum_{k=0}^{N} C_k^N x^k (1-x)^{N-k}|f(x)-f(x_k)|$$

$$= \sum\nolimits_k' + \sum\nolimits_k'' < \epsilon$$

即$|f_N(x)-f(x)| < \epsilon$ 得證。

◆Stone-Weirstrass 定理的重要性在於：任意有限區間上的連續函數都可以用多項式來“均勻地”逼近，奠定了連續函數以多項式展開的理論基礎。(注意：雖然大家早已習慣用 Taylor 展開式來表示正常函數，但 Taylor 展開先要選擇某一定點作為展開的起點，而且還需要討論「餘式」的收斂問題，故作為「理論基礎」稍為薄弱。)

補充證明

$$\sum_{k=0}^{N} C_k^N x^k (1-x)^{N-k}(x-\frac{k}{N})^2$$

$$= \frac{x(1-x)}{N}$$

由於

$$\sum_{k=0}^{N} C_k^N x^k (1-x)^{N-k} = 1$$

將上式微分，即得

$$\sum_{k=0}^{N} C_k^N x^{k-1} (1-x)^{N-k-1}(k-Nx) = 0$$

全式乘以$x(1-x)$，即為

$$\sum_{k=0}^{N} C_k^N x^k (1-x)^{N-k}(k-Nx) = 0$$

將上式再微分一次，即得

$$\sum_{k=0}^{N} C_K^N \{x^{k-1}(1-x)^{N-k-1}(k-Nx)^2$$

$$+ x^k(1-x)^{N-k}(-N)\}$$

$$= 0$$

故

$$\sum_{k=0}^{N} C_K^N x^{k-1}(1-x)^{N-k-1}(k-Nx)^2$$

$$= (\sum_{k=0}^{N} C_K^N x^k (1-x)^{N-k})\, N = N$$

$$\therefore \sum_{k=0}^{N} C_K^N x^k (1-x)^{N-k} = 1$$

將上式左右兩側皆乘以$x(1-x)$，即得

$$\sum_{k=0}^{N} C_K^N x^k (1-x)^{N-k} (k-Nx)^2$$

$$= x(1-x)N$$

即

$$\sum_{k=0}^{N} C_K^N x^k (1-x)^{N-k} (x-\frac{k}{N})^2$$

$$= \frac{x(1-x)}{N} \text{ 得證}$$

§討論

以上論證都限於[0,1]這區間上的連續函數，正如前文中所述，所得到的結論都很容易推廣到$[-L,L]$這類有限區間內，無論$L(>0)$有多大，只要L是個有限數，若要再推廣到$(-\infty,\infty)$，即$L\to\infty$的函數空間，並容納一些“不連續”但可以積分，即$\int_{-\infty}^{\infty}|f(x)|^2\,dx$存在的函數，這就需要更嚴謹的數學方法，通常被稱為 Hilbert 空間的函數即算子(operator)的理論，讀者可以自行鑽研，例如 John von Neumann 所著之“Mathematical Foundation of Quantum Mechanics”

又及：若$\int_{-\infty}^{\infty}|f(x)|^2\,dx$不收斂時，常可引入積分的權重函數(weight function)，用$W(x)$表示，

$W(x)$必須滿足：

(i)$W(x)\geq 0$

(ii)$\int_{-\infty}^{\infty} W(x)\,dx = 1$

(iii) $W(x)\to 0\ as\ x\to\pm\infty$，而且還要收斂得很快

如本文中所列出的

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t-\frac{\epsilon}{4}\omega^2}\,d\omega = \frac{1}{\sqrt{\pi\epsilon}}e^{-\frac{t^2}{\epsilon}}$$

左側那個因式$e^{-\frac{\epsilon}{4}\omega^2}$就是一個 weight function