Periodic Functions/Functions on a Torus/Fourier Series

Let $\{e_j : 1 \leq j \leq n\}$ be the standard basis in $\mathbb{R}^n$: We say $f : \mathbb{R}^n \to \mathbb{C}$ is 2π-periodic in each variable if

$$f(x + 2\pi e_j) = f(x) \qquad \forall\, x \in \mathbb{R}^n, 1 \leq j \leq n.$$

We can identify 2π-periodic functions with functions on a torus. Let $S^1 = \{e^{i\theta} : \theta \in \mathbb{R}\} \subset \mathbb{C}$, and $T^n = S^1 \times \cdots \times S^1 \subset \mathbb{C}^n$. To each function $\widetilde{\phi} : T^n \to \mathbb{C}$ we can identify a 2π-periodic function $\phi : \mathbb{R}^n \to \mathbb{C}$ by $\phi(x_1, \ldots, x_n) = \widetilde{\phi}(e^{ix_1}, \ldots, e^{ix_n})$. Conversely, each 2π-periodic function $\phi : \mathbb{R}^n \to \mathbb{C}$ induces a unique $\widetilde{\phi} : T^n \to \mathbb{C}$ for which $\widetilde{\phi}(e^{ix_1}, \ldots, e^{ix_n}) = \phi(x_1, \ldots, x_n)$. If $\phi : \mathbb{R}^n \to \mathbb{C}$ is 2π-periodic, ϕ is uniquely determined by its values $\phi(x)$ for $x \in [-\pi, \pi)^n$ or for $x \in [0, 2\pi)^n$. Let $\nu_n = \frac{1}{(2\pi)^n} m_n$, where m_n is n-dimensional Lebesgue measure. Then ν_n induces a measure $\widetilde{\nu}_n$ on T^n for which

$$\int_{T^n} \widetilde{\phi}\, d\widetilde{\nu}_n = \int_{[0,2\pi]^n} \phi\, dv_n.$$

From here on, we blur the distinction between ϕ and $\widetilde{\phi}$ and between ν_n and $\widetilde{\nu}_n$, and we will abuse these notations. Note: $\nu_n(T^n) = \nu_n([0, 2\pi]^n) = 1$. Let $L^p(T^n)$ denote $L^p([0, 2\pi]^n)$ with measure ν_n ($1 \leq p \leq \infty$). $L^2(T^n)$ is a Hilbert space with inner product

$$(\phi, \psi) = \int_{T^n} \phi\bar{\psi}\, d\nu_n = \int_{[0,2\pi]^n} \phi\bar{\psi}\, d\nu_n.$$

Theorem. $\{e^{ix\cdot\xi} : \xi \in \mathbb{Z}^n\}$ is an orthonormal system in $L^2(T^n)$.

Proof. $(e^{ix\cdot\xi}, e^{ix\cdot\eta}) = \int_{[0,2\pi]^n} e^{ix\cdot(\xi-\eta)} d\nu_n = \left\{ \begin{array}{ll} 1, & \xi = \eta \\ 0, & \xi \neq \eta \end{array} \right.$. □

Definition. A *trigonometric polynomial* is a finite linear combination of $\{e^{ix\cdot\xi} : \xi \in \mathbb{Z}^n\}$. (Note: since $\{e^{ix\cdot\xi}, e^{-ix\cdot\xi}\}$ and $\{\cos(x\cdot\xi), \sin(x\cdot\xi)\}$ span the same two-dimensional subspace, we could use sines and cosines as our basis functions.)

Definition. $C(T^n)$ is the space of all continuous 2π-periodic functions $\phi : \mathbb{R}^n \to \mathbb{C}$. Note that $C(T^n) \subsetneqq C([0, 2\pi]^n)$.

We will use the uniform norm $\|\phi\|_u = \sup_x |\phi(x)|$ on $C(T^n)$. $C^k(T^n)$ (for $k \geq 0, k \in \mathbb{Z}$) is the space of all C^k 2π-periodic functions $\phi : \mathbb{R}^n \to \mathbb{C}$. Again $C^k(T^n) \subsetneqq C^k([0, 2\pi]^n)$. We will use the norm $\|\phi\|_{C^k} = \sum_{|\alpha|\leq k} \|\partial^\alpha \phi\|_u$.

Fourier Coefficients

For $f \in L^1(T^n)$, define $\widehat{f}(\xi) = \int_{T^n} e^{-ix\cdot\xi} f(x)d\nu_n(x)$ for $\xi \in \mathbb{Z}^n$. Then for $f \in L^2(T^n)$, $\widehat{f}(\xi) = (f, e^{ix\cdot\xi})$. (Note that $\|f\|_1 \le \|f\|_2$ since $\nu(T^n) = 1$.)

The *Fourier series* of f is the *formal series* $\sum_{\xi\in\mathbb{Z}^n} \widehat{f}(\xi)e^{ix\cdot\xi}$. We will study in what sense this series converges to f.

Lemma. Let $a_k = \int_{T^n} \left(\prod_{j=1}^n \frac{1+\cos x_j}{2}\right)^k d\nu_n(x)$ (note: $0 \le$ integrand ≤ 1). Then $a_k \ge \left(\frac{2}{\pi(k+1)}\right)^n$. Define $q_k(x) = \frac{1}{a_k}\left(\prod_{j=1}^n \frac{1+\cos x_j}{2}\right)^k$. Then

(1) q_k is a trigonometric polynomial (of degree nk)

(2) $q_k(x) \ge 0$

(3) $\int_{T^n} q_k(x)d\nu_n(x) = 1$

(4) If $\eta_k(\delta) = \max\{q_k(x) : x \in [-\pi,\pi]^n \backslash (-\delta,\delta)^n\}$ then $\lim_{k\to\infty} \eta_k(\delta) = 0$ (for such x, $q_k(x) \le \frac{1}{a_k}\left(\frac{1+\cos\delta}{2}\right)^k \le \left(\frac{\pi(k+1)}{2}\right)^2 \left(\frac{1+\cos\delta}{2}\right)^k \to 0$ as $k \to \infty$).

Theorem. Given $f \in C(T^n)$, let

$$p_k(x) = (f * q_k)(x) = \int_{T^n} f(x-y)q_k(y)d\nu_n(y).$$

Then p_k is a trig polynomial, and $\|p_k - f\|_u \to 0$ as $k \to \infty$.

Proof. Since $\widehat{q}_k(\xi) = 0$ for $|\xi|$ sufficiently large,

$$\begin{aligned} p_k(x) &= \int_{T^n} f(y)q_k(x-y)d\nu_n(y) = \int_{T^n} f(y)\sum_\xi \widehat{q}_k(\xi)e^{i(x-y)\cdot\xi}d\nu_n(y) \\ &= \sum_\xi \widehat{f}(\xi)\widehat{q}_k(\xi)e^{ix\cdot\xi} \end{aligned}$$

is a trig. poly. Given $\varepsilon > 0$, choose δ (by the uniform continuity of f) $\ni$

$$|x-w|_\infty < \delta \Rightarrow |f(x) - f(w)| < \varepsilon.$$

Then $p_k(x) - f(x) = \int_{T^n}(f(x-y) - f(x))q_k(y)d\nu_n(y)$

$$|p_k(x) - f(x)| \le \int_{T^n} |f(x-y) - f(x)|q_k(y)d\nu_n(y) = \underbrace{I_1}_{(-\delta,\delta)^n} + \underbrace{I_2}_{[-\pi,\pi]^n\backslash(-\delta,\delta)^n}$$

$$\begin{aligned} I_1 &\le \int_{(-\delta,\delta)^n} \varepsilon q_k(y)d\nu_n(y) \le \varepsilon \\ I_2 &\le \int_{[-\pi,\pi]^n\backslash(-\delta,\delta)^n} 2\|f\|_u\eta_k(\delta)d\nu_n \le 2\|f\|_u\eta_k(\delta) < \varepsilon \end{aligned}$$

for k suff. large. $\square$

Corollary 1 *Trig polynomials are dense in $C(T^n)$.*

Remark. Sequences q_k with properties (1), (2), (3), (4) are called *summability kernels*. In $\mathbb{R}^1$, another such kernel is the Féjer kernel $q_k(x) = \frac{1}{k+1}\frac{\sin^2\left(\frac{k+1}{2}x\right)}{\sin^2\left(\frac{1}{2}x\right)} = \sum_{\xi=-k}^{k}\left(1-\frac{|\xi|}{k+1}\right)e^{ix\xi}$. If we define $S_k(f) = \sum_{\xi=-k}^{k}\widehat{f}(\xi)e^{ix\cdot\xi}$ and $\sigma_k(f) = \frac{1}{k+1}(S_0(f)+\cdots+S_k(f)) = f * q_k$ then for $f \in C(T)$, $\sigma_k(f) \to f$ uniformly (same proof as above). (This is the classical result that the Fourier series of an $f \in C(T)$ is Cesàro summable to f.)

Corollary 2 *Trig polynomials are dense in $L^2(T^n)$.*

Proof. Given $f \in L^2(T^n)$ and $\varepsilon > 0$, $\exists\, g \in C(T^n) \ni \|f-g\|_2 < \frac{\varepsilon}{2}$. $\exists$ trig. poly $p \ni \|p-g\|_u < \frac{\varepsilon}{2}$, so since $\nu_n(T^n) = 1$,

$$\|f-p\|_2 \le \|f-g\|_2 + \|g-p\|_2 \le \|f-g\|_2 + \|g-p\|_u < \varepsilon.$$

$\square$

Corollary 3 $\{e^{ix\cdot\xi} : \xi \in \mathbb{Z}^n\}$ *is a complete orthonormal system in $L^2(T^n)$. Hence if $f \in L^2(T^n)$, the Fourier series of f (any arrangement) converges to f in L^2. Also, the map $\mathcal{F} : L^2(T^n) \to l^2(\mathbb{Z}^n)$ given by $f \mapsto \widehat{f}$ is a Hilbert space isomorphism.*

Theorem. If $f \in L^1(T^n)$, then $p_k \to f$ in $L^1(T^n)$ (where $p_k = f * q_k$ and q_k is given on the previous page).

Proof. The proof is similar to the proof of the Theorem above, except we use continuity of translation in L^1 instead of uniform continuity. Given $\varepsilon > 0$, choose $\delta \ni \|f(x-\alpha)-f(x)\|_1 < \varepsilon$ whenever $|\alpha|_\infty < \delta$. By Fubini,

$$\begin{aligned}\int_{T^n}|p_k(x)-f(x)|d\nu_n(x) &\le \int_{T^n}\left[q_k(y)\int_{T^n}|f(x-y)-f(x)|d\nu_n(x)\right]d\nu_n(y)\\ &= I_1 \rightarrow (-\delta,\delta)^n + I_2 \rightarrow [-\pi,\pi]^n\backslash(-\delta,\delta)^n\end{aligned}$$

$$\begin{aligned}I_1 &\le \textstyle\int q_k(y)\varepsilon d\nu_n(y) = \varepsilon\\ I_2 &\le 2\|f\|_1\eta_k(\delta) \to 0 \text{ as } k\to\infty\end{aligned}$$

$\square$

Corollary. (Uniqueness Theorem). If $f \in L^1(T^n)$ and $(\forall\,\xi \in \mathbb{Z}^n)\widehat{f}(\xi) = 0$, then $f = 0$ a.e. (Thus if $f, g \in L^1(T^n)$ and $\widehat{f} \equiv \widehat{g}$, then $f = g$ a.e.)

Proof. If $\widehat{f} \equiv 0$, then $p_k(x) = \sum_\xi \widehat{f}(\xi)\widehat{q_k}(\xi)e^{ix\cdot\xi} = 0$, and $p_k \to f$ in L^1. $\square$

Theorem. (Riemann-Lebesgue Lemma). If $f \in L^1(T^n)$, then $\widehat{f}(\xi) \to 0$ as $|\xi| \to \infty$.

Proof. View f as $f(x)\chi_{[-\pi,\pi]^n}(x) \in L^1(\mathbb{R}^n)$. $\square$

Absolutely Convergent Fourier Series

Suppose $f \in L^1(T^n)$ and $\widehat{f} \in l^1(\mathbb{Z}^n)$. Then the Fourier series of f converges absolutely and uniformly to a $g \in C(T^n)$, and $g = f$ a.e.

Proof. Let $g(x) = \sum_{\xi \in \mathbb{Z}^n} \widehat{f}(\xi) e^{ix\cdot\xi}$ (converges uniformly and absolutely). Then $g \in C(T^n)$. By the Dominated Convergence Theorem,

$$\begin{aligned} \widehat{g}(\xi) &= \int_{T^n} e^{-ix\cdot\xi} \left(\sum_{\eta\in\mathbb{Z}^n} \widehat{f}(\eta) e^{ix\cdot\eta} \right) d\nu_n(x) \\ &= \sum_{\eta\in\mathbb{Z}^n} \widehat{f}(\eta) \int_{T^n} e^{-ix\cdot\xi} e^{ix\cdot\eta} d\nu_n(x) = \widehat{f}(\xi). \end{aligned}$$

So $g = f$ a.e. $\square$

Decay of Fourier Coefficients $\leftrightarrow$ Smoothness of f

Lemma. Suppose $\alpha(\xi) \in l^1(\mathbb{Z}^n)$ and $(i\xi_j)\alpha(\xi) \in l^1(\mathbb{Z}^n)$. Let $f = \sum_\xi \alpha(\xi) e^{ix\cdot\xi}$, $g = \sum_\xi (i\xi_j)\alpha(\xi) e^{ix\cdot\xi}$. Then $f, g \in C(T)$, $\frac{\partial f}{\partial x_j}$ exists everywhere, and $\frac{\partial f}{\partial x_j} = g$.

Proof. The two series of continuous functions converge absolutely and uniformly to f and g, respectively. Since $\frac{\partial}{\partial x_j}(\alpha(\xi)e^{ix\cdot\xi}) = i\xi_j \alpha(\xi) e^{ix\cdot\xi}$, the result follows from a standard theorem in analysis. $\square$

Theorem. Suppose $f \in L^1(T^n)$ and $(1 + |\xi|^m)\widehat{f}(\xi) \in l^1(\mathbb{Z}^n)$ for some integer $m \geq 0$. Then the Fourier series of f converges absolutely and uniformly to a $g \in C^m(T^n)$, and $f = g$ a.e.

Proof. Just must show $g \in C^m(T^n)$. For each ν with $|\nu| \leq m$, $(i\xi)^\nu \widetilde{f}(\xi) \in l^1(\mathbb{Z}^n)$, so $\sum_\xi (i\xi)^\nu \widetilde{f}(\xi) e^{ix\cdot\xi}$ converges abs. unif. to some $g_\nu \in C(T^n)$. By the Lemma and induction, $g_\nu = \partial^\nu g$. $\square$

Theorem. Suppose $f \in C^m(T^n)$.

(a) For $|\nu| \leq m$, $\widehat{\partial_x^\nu f}(\xi) = (i\xi)^\nu \widehat{f}(\xi)$

(b) So $(1 + |\xi|^2)^{\frac{m}{2}} \widehat{f}(\xi)$ (or $(1 + |\xi|^m)\widehat{f}(\xi)$ or $(1 + |\xi|)^m \widehat{f}(\xi)$) $\in l^2(\mathbb{Z}^n)$

(c) Hence if $k < m - \frac{n}{2}$, then $(1 + |\xi|^2)^{\frac{k}{2}} |\widehat{f}(\xi)| \in l^1(\mathbb{Z}^n)$ and $\exists\, C_k \ni |\widehat{f}(\xi)| \leq C_k |\xi|^{-k}$.

Proof.

(a) Integration by parts — e.g., let $\widetilde{x} = (x_2, \ldots, x_n)$, so $x = (x_1, \widetilde{x})$

$$\begin{aligned} \widehat{\frac{\partial f}{\partial x_1}}(\xi) &= \int_{T^{n-1}} d\nu_{n-1}(\widetilde{x}) \int_T d\nu_1(x_1) e^{-ix\cdot\xi} \frac{\partial f}{\partial x_1}(x) \\ &= \int_{T^{n-1}} d\nu_{n-1}(\widetilde{x})(i\xi_1) \int_T d\nu_1(x_1) e^{-ix\cdot\xi} f(x) \\ &= (i\xi_1)\widehat{f}(\xi). \end{aligned}$$

(b) Choosing $\nu = 0$ and $\nu = me_j$ for $j = 1, \ldots, n$, we get $(1 + |\xi_1|^m + \cdots + |\xi_n|^m)|\widehat{f}(\xi)| \in l^2(\mathbb{Z}^n)$, (b) follows.

(c) Suppose $k < m - \frac{n}{2}$. By Cauchy-Schwarz, $\left(\frac{k}{2} = \frac{k-m}{2} + \frac{m}{2}\right)$

$$\sum_{\xi}(1+|\xi|^2)^{\frac{k}{2}}|\widehat{f}(\xi)| \le \underbrace{\left(\sum(1+|\xi|^2)^{k-m}\right)^{\frac{1}{2}}}_{<\infty \text{ since } 2(k-m)<-n} \underbrace{\left(\sum(1+|\xi|^2)^m|\widehat{f}(\xi)|^2\right)^{\frac{1}{2}}}_{<\infty} \equiv C_k.$$

For each $\xi \in Z^n$, $(1+|\xi|^2)^{\frac{k}{2}}|\widehat{f}(\xi)| \le C_k$, so $|\widehat{f}(\xi)| \le C_k|\xi|^{-k}$.

□

The Hausdorff-Young Inequality. Suppose $1 \le p \le 2$ and $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(T^n)$, then $\widehat{f} \in l^q(\mathbb{Z}^n)$, and $\|\widehat{f}\|_q \le \|f\|_p$.

Proof. For $f \in L^1(T^n)$, $\|\widehat{f}\|_\infty \le \|f\|_1$ and for $f \in L^2(T^n)$, $\|\widehat{f}\|_2 = \|f\|_2$. By the Riesz-Thorin interpolation theorem, $\|\widehat{f}\|_q \le \|f\|_p$. □

Other decay estimates for $f \in L^1(T)$. (See Katznelson, *Intro to Harmonic Analysis*, pp. 24–25.)

(1) If $f \in L^1(T)$, then $\widehat{f}(\xi) = o(1)$ (i.e., $\widehat{f}(\xi) \to 0$).

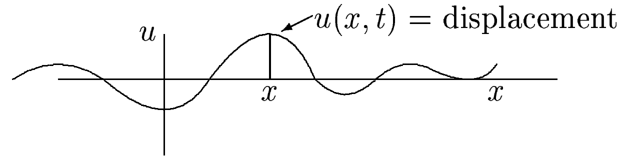
(2) If f is absolutely continuous (i.e, $f' \in L^1$, etc.), then $\widehat{f}(\xi) = o\left(\frac{1}{|\xi|}\right)$ (i.e., $|\xi|\widehat{f}(\xi) \to 0$).

(3) If $f \in BV(T)$, then $\widehat{f}(\xi) = \mathcal{O}\left(\frac{1}{|\xi|}\right)$ (in fact $|\widehat{f}(\xi)| \le \frac{\mathrm{Var}(f)}{2\pi|\xi|}$).

Application: Vibrating Strings

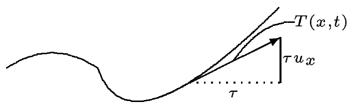
Consider an infinite oscillating string. Assume that the x-axis is the equilibrium position of the string and that the tension in the string at rest in equilibrium is τ. Let $u(x,t)$ denote the displacement at x at time t. Then the *wave equation* (in one space dimension) governs the motion.

"snap shot" at time t:

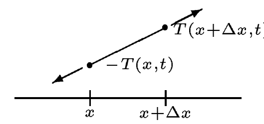


Derivation (for small displacements). We make the following simplifying assumptions:

- the displacement of the string from equilibrium (and its slope) are small;
- each point on the string moves only in the vertical direction;
- the tension force $T(x,t)$ in the string (i.e., the (vector) force which the part of the string to the right of x exerts on the part to the left of x, at time t) is tangential to the string and has magnitude proportional to the local stretching factor $\sqrt{1+u_x^2}$.



Since $u_x = 0$ in equilibrium, the constant of proportionality is the equilibrium tension τ. Thus the magnitude of $T(x,t)$ is $\tau\sqrt{1+u_x(x,t)^2}$, and the vertical component of $T(x,t)$ is τu_x. Now consider the part of the string between x and $x+\Delta x$. The vertical component of Newton's second law ($F = ma$, force $=$ mass $\cdot$ acceleration) applied to this part of the string is



$$\overbrace{\tau u_x(x+\Delta x,t) - \tau u_x(x,t)}^{\text{force}} = \overbrace{\rho\Delta x}^{\text{mass}}\,\overbrace{u_{tt}(x,t)}^{\text{accel}},$$

where ρ is the density (mass per unit length; assumed constant). Dividing by Δx and taking the limit as $\Delta x \to 0$, we obtain

$$\tau u_{xx} = \rho u_{tt}.$$

Normalizing units so that $\rho = \tau$, we obtain the *wave equation* (in one space dimension):

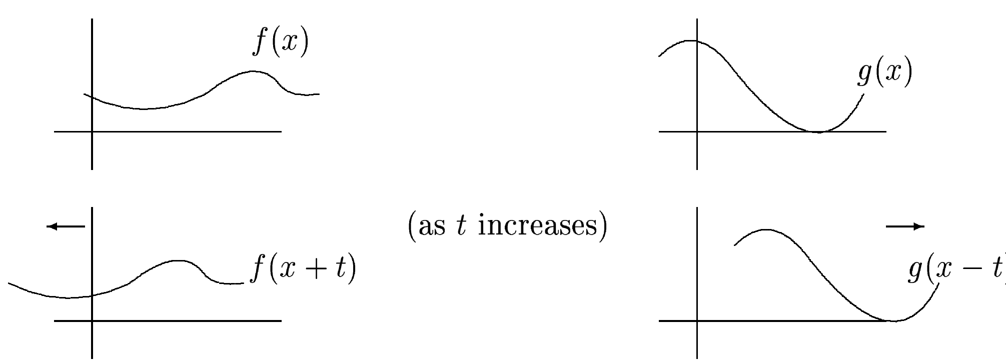
$$u_{tt} = u_{xx}.$$

Solutions of $u_{tt} = u_{xx}$

Change variables. Let $y = x+t$, $z = x-t$ (so $x = \frac{y+z}{2}$, $t = \frac{y-z}{2}$). Then $\frac{\partial}{\partial y} = \frac{\partial x}{\partial y}\frac{\partial}{\partial x} + \frac{\partial t}{\partial y}\frac{\partial}{\partial t} = \frac{1}{2}\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right)$ and $\frac{\partial}{\partial z} = \frac{\partial x}{\partial z}\frac{\partial}{\partial x} + \frac{\partial t}{\partial z}\frac{\partial}{\partial t} = \frac{1}{2}\left(\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\right)$, so $u_y = \frac{1}{2}(u_x + u_t)$ and $u_{yz} = \frac{1}{2}\left(\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\right)\frac{1}{2}(u_x+u_t) = \frac{1}{4}(u_{xx} - u_{tt})$. In the new coordinates, the wave equation becomes simply $u_{yz} = 0$. Thus u_y is independent of z, i.e., $u_y = \widetilde{f}(y)$. Integrating in y for each fixed z, we get $u = f(y) + g(z)$ (where $f(y) = \int \widetilde{f}(y)dy$). So any solution of the wave equation $u_{tt} = u_{xx}$ is of the form

$$u(x,t) = f(x+t) + g(x-t). \tag{$*$}$$

Physically, this is a superposition of left-going and right-going waves:



Observation. The derivation above shows that any C^2 function of x and t satisfying the wave equation is of the form ($*$). Conversely, if f and g are C^2 functions of one variable, it is easily checked that $u(x,t) = f(x+t) + g(x-t)$ is a C^2 solution of the wave equation. But if f and g are only continuous, $f(x+t) + g(x-t)$ still makes sense; in what sense is this a solution of $u_{tt} = u_{xx}$? (It is true in the sense of distributions.)

Initial-Value Problem (IVP) (or the *Cauchy Problem*). Thinking only in terms of ODEs in time t (overlooking that $\frac{\partial^2}{\partial x^2}$ is not a bounded linear operator) we "should" be able to determine $u(x,t)$ for $x \in \mathbb{R}$ and $t \geq 0$ if we are given initial values $u(x,0)$ and $u_t(x,0)$ for $x \in \mathbb{R}$ (we need u and u_t at $t = 0$ since the equation is second-order in t).

D'Alembert's Formula (for the Cauchy Problem for $u_{tt} = u_{xx}$). Consider the IVP: DE $u_{tt} = u_{xx}$ $(x \in \mathbb{R}, t \geq 0)$

$$\text{IC} \left\{ \begin{array}{rcll} u(x,0) & = & f(x) & (x \in \mathbb{R}) \\ u_t(x,0) & = & g(x) & (x \in \mathbb{R}) \end{array} \right. \qquad \text{(note: not the same } f,\ g \text{ as above)}$$

(To obtain a C^2 solution $u(x,t)$, it will suffice for $f \in C^2(\mathbb{R})$, $g \in C^1(\mathbb{R})$.) We will separately analyze the cases $g \equiv 0$ and $f \equiv 0$, and then use superposition.

Case 1. $g \equiv 0$ IC $\left\{ \begin{array}{rcl} u(x,0) & = & f(x) \\ u_t(x,0) & = & 0 \end{array} \right.$ $(x \in \mathbb{R})$. We have $u(x,t) = w_1(x+t) + w_2(x-t)$ for some $w_1, w_2 \in C^2(\mathbb{R})$. By the IC, $\begin{array}{rcl} w_1(x) + w_2(x) & = & u(x,0) = f(x) \\ w_1'(x) - w_2'(x) & = & u_t(x,0) = 0 \end{array}$ (so w_1 and w_2 differ by a constant). One solution is $w_1(x) = w_2(x) = \frac{1}{2}f(x)$.

Remark. For a solution $u(x,t)$ of $u_{tt} = u_{xx}$, w_1 and w_2 are uniquely determined up to a constant (if $w_1(x+t) + w_2(x-t) = v_1(x+t) + v_2(x-t)$, then $w_1(x+t) - v_1(x+t) = v_2(x-t) - w_2(x-t)$ is independent of both $y = x+t$ and $z = x-t$, and is thus a constant).

So any other solution is $\left\{ \begin{array}{rcl} w_1(x) & = & \frac{1}{2}f(x) + c \\ w_2(x) & = & \frac{1}{2}f(x) - c \end{array} \right.$ for some constant c. So the solution to Case 1 is

$$u(x,t) = \frac{1}{2}f(x+t) + \frac{1}{2}f(x-t).$$

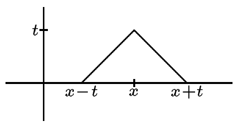
Case 2. $f \equiv 0$ IC $\left\{ \begin{array}{rcl} u(x,0) & = & 0 \\ u_t(x,0) & = & g(x) \end{array} \right.$ $(x \in \mathbb{R})$. Again, $u(x,t) = w_1(x+t) + w_2(x-t)$ for some $w_1, w_2 \in C^2(\mathbb{R})$. By the IC, $\left\{ \begin{array}{rcl} w_1(x) + w_2(x) & = & 0 \\ w_1'(x) - w_2'(x) & = & g(x) \end{array} \right. \Rightarrow \left\{ \begin{array}{rcl} w_2 & = & -w_1 \\ w_1' & = & \frac{1}{2}g \end{array} \right. \Rightarrow w_1 = \frac{1}{2}\int g$. So the solution to Case 2 is

$$u(x,t) = \left(\frac{1}{2}\int g \right) \Bigg|_{x-t}^{x+t} = \frac{1}{2}\int_{x-t}^{x+t} g(s)ds.$$

Adding Cases 1 and 2, the solution of the IVP with IC $\left\{ \begin{array}{rcl} u(x,0) & = & f(x) \\ u_t(x,0) & = & g(x) \end{array} \right.$ is:

$$u(x,t) = \frac{1}{2}f(x+t) + \frac{1}{2}f(x-t) + \frac{1}{2}\int_{x-t}^{x+t} g(s)ds \qquad \text{d'Alembert's formula}$$

Remarks.

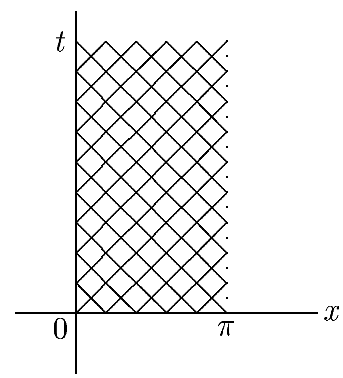


(1) d'Alembert's formula gives an explicit demonstration of the *finite domain of dependence* of the solution of this IVP on the initial data (a general property of hyperbolic PDEs): for a fixed $x \in \mathbb{R}$ and fixed $t > 0$, $u(x,t)$ depends only on $f(x+t)$, $f(x-t)$, and $\{g(s) : x-t \le s \le x+t\}$.

(2) d'Alembert's formula also provides a solution for negative t as well: $u_{tt} = u_{xx}$ $(x \in \mathbb{R}, t \le 0)$, $\begin{cases} u(x,0) &= f(x) \\ u_t(x,0) &= g(x) \end{cases}$ ("final" conditions); like ODEs, hyperbolic PDEs in general can be advanced either in the $+t$ direction or the $-t$ direction.

Initial-Boundary Value Problem (IBVP)

Consider now a finite string $(0 \le x \le \pi)$ fixed at both ends, so $u(0,t) = u(\pi,t) \equiv 0$. Suppose the initial displacement is $u(x,0) = f(x)$ $(0 \le x \le \pi)$ (where $f(0) = f(\pi) = 0$), and for simplicity suppose the initial velocity is $u_t(x,0) = 0$ $(0 \le x \le \pi)$. This models a "plucked" violin string (moved to position $u(x,0) = f(x)$ at time $t = 0$, and then released with initial velocity $u_t(x,0) = 0$). We obtain an IBVP with both initial conditions (IC) and boundary conditions (BC):



$$\begin{array}{lll} \text{DE} & u_{tt} = u_{xx} & (0 \le x \le \pi, t \ge 0) \\ \text{IC} & \begin{cases} u(x,0) &= f(x) \\ u_t(x,0) &= 0 \end{cases} & (0 \le x \le \pi) \\ \text{BC} & \begin{cases} u(0,t) &= 0 \\ u(\pi,t) &= 0 \end{cases} & (t \ge 0) \end{array}$$

We will solve this IBVP in two ways: ① by d'Alembert's formula, and ② by Fourier series.

Solution ① (d'Alembert). Find functions w_1, w_2 defined on $\mathbb{R}$ so that $u(x,t) = w_1(x+t) + w_2(x-t)$ satisfies the IC and BC. The BC $u(0,t) = 0$ for $t \ge 0$ gives $0 = w_1(t) + w_2(-t)$ for $t \ge 0$, or $w_2(t) = -w_1(-t)$ for $t \le 0$. [Note that to define $u(x,t)$ in the region $0 \le x \le \pi$, $t \ge 0$, we only need to give $w_1(s)$ for $s \ge 0$ and $w_2(s)$ for $s \le \pi$. To simplify our calculations, we will find w_1 and w_2 defined on all of $\mathbb{R}$, so that $u(x,t)$ satisfies the BC for $t \le 0$ too.] So we ask $w_2(t) = -w_1(-t) (\forall\, t \in \mathbb{R})$. Next, the BC $u(\pi,t) = 0$ (now $\forall\, t \in \mathbb{R}$) gives $0 = w_1(\pi+t) + w_2(\pi-t)$, so $w_1(\pi+t) = -w_2(\pi-t) = w_1(t-\pi)$, i.e., $w_1(t+2\pi) = w_1(t) (\forall\, t \in \mathbb{R})$. So w_1 is 2π-periodic, and thus $w_2(t) = -w_1(-t)$ is also 2π-periodic. The IC $u_t(x,0) = 0$ $(0 \le x \le \pi)$ gives $0 = w_1'(x) - w_2'(x) = w_1'(x) - w_1'(-x)$ for $0 \le x \le \pi$, i.e., $w_1'(-x) = w_1'(x)$ for $0 \le x \le \pi$; since w_1' is 2π-periodic, we conclude that w_1' is an even function on $\mathbb{R}$. We may assume $w_1(0) = 0$ (if not, replace w_1 by $w_1(s) - w_1(0)$ and replace w_2 by $w_2(s) + w_1(0)$). Then $w_1(-x) = \int_0^{-x} w_1'(s)ds = -\int_0^x w_1'(-s)ds = -\int_0^x w_1'(s)ds = -w_1(x) (\forall\, x \in \mathbb{R})$, so

w_1 is an odd function on $\mathbb{R}$; moreover $w_2 = w_1$ since $w_2(t) = -w_1(-t)$. Finally, the IC $u(x,0) = f(x)$ $(0 \le x \le \pi)$ gives $f(x) = w_1(x) + w_2(x) = 2w_1(x)$, i.e., $w_1(x) = \frac{1}{2}f(x)$ for $0 \le x \le \pi$. This completes the determination of w_1: it is the 2π-periodic, odd function on $\mathbb{R}$ which agrees with $\frac{1}{2}f$ on $[0,\pi]$. So d'Alembert's solution can be summarized as follows: define $\widetilde{f}(x) = f(x)$ for $0 \le x \le \pi$, $\widetilde{f}(x) = -f(-x)$ for $-\pi \le x \le 0$ (the odd extension of f from $[0,\pi]$ to $[-\pi,\pi]$), and then extend $\widetilde{f}$ to be 2π-periodic on $\mathbb{R}$. [Note: if $f(0) = f(\pi) = 0$ and $f \in C^1[0,\pi]$, then $\widetilde{f} \in C^1(\mathbb{R})$; if in addition $f \in C^2[0,\pi]$ and $f''(0) = f''(\pi) = 0$, then $\widetilde{f} \in C^2(\mathbb{R})$.] We obtain d'Alembert's formula for the solution of this IBVP:

$$u(x,t) = \frac{1}{2}\left(\widetilde{f}(x+t) + \widetilde{f}(x-t)\right)$$

(remember, this is the special case where $u_t(x,0) = 0$ $(0 \le x \le \pi)$).

Solution ② (Fourier series). We use separation of variables. We want to find simple harmonics of the string, that is, solutions of the form $u(x,t) = v(x)w(t)$ (often called *fundamental modes*). Using $'$ to mean $\frac{d}{dx}$ for v, and also $\frac{d}{dt}$ for w, the DE $u_{tt} = u_{xx}$ becomes $v(x)w''(t) = v''(x)w(t)$, or (wherever $v(x)w(t) \ne 0$) $\frac{w''(t)}{w(t)} = \frac{v''(x)}{v(x)}$. The LHS is independent of x and the RHS is independent of t, so both sides are equal to a constant; call it $-\lambda$.

We end up with ODEs for v and w:

$$\begin{array}{rclll} v''(x) + \lambda v(x) & = & 0 & (0 \le x \le \pi) & \text{"spatial ODE"} \\ w''(t) + \lambda w(t) & = & 0 & (t \ge 0 & \text{"temporal ODE"} \end{array}$$

Applying the BC to the "spatial ODE", we get $v(0) = v(\pi) = 0$, leading to the following "eigenvalue problem:" determine for which (in this case real) values of λ there exists a non-trivial (i.e., not $\equiv 0$) solution $v(x)$ of the boundary-value problem (BVP):

$$\begin{array}{lrcll} \text{DE} & v'' + \lambda v & = & 0 & 0 \le x \le \pi \\ \text{BC} & v(0) = v(\pi) & = & 0 & \end{array}$$

Case (i) $\lambda < 0$: The general solution of $v'' + \lambda v = 0$ is $c_1 \cosh(\sqrt{-\lambda}x) + c_2 \sinh(\sqrt{-\lambda}x)$. $v(0) = 0 \Rightarrow c_1 = 0$, and then $v(\pi) = 0 \Rightarrow c_2 = 0$. No nontrivial solutions.

Case (ii) $\lambda = 0$: The general solution of $v'' = 0$ is $v(x) = c_1 + c_2x$. $v(0) = 0 \Rightarrow c_1 = 0$, and then $v(\pi) = 0 \Rightarrow c_2 = 0$. No nontrivial solutions.

Case (iii) $\lambda > 0$: The general solution of $v'' + \lambda v = 0$ is $v(x) = c_1 \cos(\sqrt{\lambda}x) + c_2 \sin(\sqrt{\lambda}x)$ $v(0) = 0 \Rightarrow c_1 = 0$. Then $v(\pi) = 0$ (and $c_2 \ne 0$ so v is nontrivial) $\Rightarrow \sin(\sqrt{\lambda}\pi) = 0 \Rightarrow \sqrt{\lambda} \in \{1,2,3,\ldots\} \Rightarrow \lambda = n^2$ for $n \in \{1,2,3,\ldots\}$. These are the "eigenvalues" of this eigenvalue problem. The corresponding "eigenfunctions" are $\sin(\sqrt{\lambda}) = \sin(nx)$.

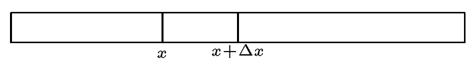
Applying the homogeneous IC $u_t(x,0) = 0$ to the "temporal ODE," we get $w'(0) = 0$. For $\lambda = n^2$, the general solution of $w'' + \lambda w = 0$ is $c_1 \cos nt + c_2 \sin nt$. The IC $w'(0) = 0$ implies $c_2 = 0$, so $w(t) = c_1 \cos nt$. Thus the *fundamental modes* for this problem are

$$u_n(x,t) = \cos(nt)\sin(nx) \qquad n \in \{1,2,3,\ldots\}.$$

Linear combinations of these are also solutions of the DE, the BC, and the one IC $u_t(x,0) = 0$. To satisfy the IC $u(x,0) = f(x)$ for $0 \le x \le \pi$, we represent $f(x)$ in a Fourier sine series: $f(x) = \sum_{n=1}^{\infty} A_n \sin(nx)$. Then (provided this series converges appropriately) $u(x,t) = \sum_{n=1}^{\infty} A_n \cos(nt)\sin(nx)$ satisfies the DE, the BC, and both IC. (See Problem 3 on Problem Set 7 for details).

Application: Heat Flow

Consider heat flow in a thin rod with insulated lateral surface.



Assume that the temperature $u(x,t)$ is a function only of horizontal position x and time t. By Newton's law of cooling, the amount of heat flowing from left to right across the point x in time Δt is $-\kappa \frac{\partial u}{\partial x}(x,t)\Delta t$ (proportional to the gradient of temperature), where the constant of proportionality κ is called the *heat conductivity* of the rod. So the net heat flowing *into* the part of rod between x an $x + \Delta x$ in the time interval from t to $t + \Delta t$ is

$$\kappa \frac{\partial u}{\partial x}(x + \Delta x, t)\Delta t - \kappa \frac{\partial u}{\partial x}(x,t)\Delta t.$$

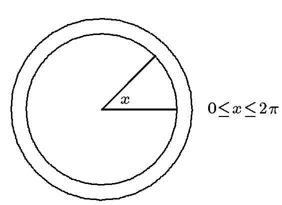
The net heat flowing *into* this part of the rod in this time interval can also be expressed as $\overbrace{\rho\Delta x}^{\text{mass}} \cdot \overbrace{c}^{\text{specificheat}} \cdot \overbrace{\frac{\partial u}{\partial t}\Delta t}^{\approx \Delta u}$, where ρ is the density (mass per unit length) of the rod, and c is the *specific heat* of the rod (the amount of heat needed to raise a unit mass by 1 unit of temperature). Equating these two expressions, dividing by Δt and Δx, and taking the limit as $\Delta x \to 0$, we obtain

$$\kappa u_{xx} = \rho c u_t.$$

Normalizing units so that $\rho c = \kappa$, we obtain the *heat equation* (in one space dimension):

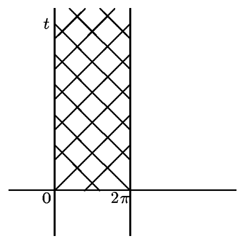
$$u_t = u_{xx}.$$

Fourier considered circular rods of length 2π, leading to the following IBVP with periodic BC:



IBVP:

	DE	$u_t = u_{xx}$	$0 \le x \le 2\pi, t \ge 0$
	IC	$u(x,0) = f(x)$	$0 \le x \le 2\pi$
	periodic BC	$\begin{cases} u(0,t) = u(2\pi,t) \\ u_x(0,t) = u_x(2\pi,t) \end{cases}$	$t \ge 0$



(We can view u defined on $T \times [0,\infty)$ [where $T = S^1$], or 2π-periodic func. of $x \in \mathbb{R}$ with $t \ge 0$.)

As with the wave equation, we separate variables, and look for solutions of the form $u(x,t) = v(x)w(t)$. The DE $u_t = u_{xx}$ becomes $v(x)w'(t) = v''(x)w(t)$, or (wherever $v(x)w(t) \ne 0$) $\frac{w'}{w} = \frac{v''}{v}$; both sides are equal to a constant; call it $-\lambda$. The "spatial ODE" is $v''(x) + \lambda v(x) = 0$ and the "temporal ODE" is $w'(t) + \lambda w(t) = 0$ $(t \ge 0)$.

Eigenvalue Problem: $v'' + \lambda v = 0$ $(0 \le x \le 2\pi)$

$$\text{periodic BC } v(0) = v(2\pi), \qquad v'(0) = v'(2\pi).$$

Case (i). $\lambda < 0$ $v \equiv 0$
Case (ii). $\lambda = 0$ One lin. ind. solution: $v \equiv 1$
Case (iii). $\lambda > 0$ $\lambda = n^2$ for $n \in \{1,2,3,\ldots\}$, with two lin. ind. solutions: $\cos(nt)$ and $\sin(nt)$ (see Problem 1 on Problem Set 7 for details). For $\lambda = n^2$ (with $n \in \{0,1,2,\ldots\}$), one lin. ind. soln. of $w' + \lambda w = 0$: $w = e^{-\lambda t}$. Thus the *fundamental modes* for this problem are $u \equiv 1$ and for $n \in \{1,2,3,\ldots\}$, $u(x,t) = e^{-n^2 t}\cos nt$ and $u(x,t) = e^{-n^2 t}\sin nt$. To satisfy the IC $u(x,0) = f(x)$ for $0 \le x \le 2\pi$, we represent $f(x)$ in a Fourier series: $f(x) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$. Then (provided this series converges appropriately)

$$u(x,t) = a_0 + \sum_{n=1}^{\infty} e^{-n^2 t}(a_n \cos nx + b_n \sin nx)$$

satisfies the DE, the periodic BC, and the IC.

Remark. This form of the Fourier series of f (viewed as its 2π-periodic extension) is equivalent to the complex exponential form $f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$. For $n \ge 1$, $\cos nx = \frac{1}{2}(e^{inx} + e^{-inx})$ and $\sin nx = \frac{1}{2i}(e^{inx} - e^{-inx})$ span the same two-dimensional subspace (over $\mathbb{C}$) as $e^{inx} = \cos nx + i \sin nx$ and $e^{-inx} = \cos nx - i \sin nx$. The coefficients are related as follows: $c_0 = a_0$; for $n \ge 1$, $c_n = \frac{1}{2}(a_n - ib_n)$, $c_{-n} = \frac{1}{2}(a_n + ib_n)$, $a_n = c_n + c_{-n}$, $b_n = i(c_n - c_{-n})$. In the inner product $\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx$ on $L^2(T)$ (here $T = S^1$), the set $\{1\} \bigcup \{\sqrt{2}\cos nx : n \ge 1\} \bigcup \{\sqrt{2}\sin nx : n \ge 1\}$ is a complete orthonormal set in $L^2(T)$, giving us the following formulas for a_n and b_n:

$$a_0 = \langle f, 1\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(x)dx.$$

$$\text{For } n \geq 1, \quad \begin{aligned} a_n &= \frac{\langle f, \cos nx\rangle}{\langle \cos nx, \cos nx\rangle} &= 2\langle f, \cos nx\rangle &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx dx \\ b_n &= \frac{\langle f, \sin nx\rangle}{\langle \sin nx, \sin nx\rangle} &= 2\langle f, \sin nx\rangle &= \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx dx. \end{aligned}$$

Caution. Many books will write $f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$, in which case $a_0 = 2\langle f, 1\rangle = \frac{1}{\pi}\int_0^{2\pi} f(x)dx = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(0x)dx$.

The solution $u(x,t)$ expressed in terms of complex exponentials is

$$u(x,t) = \sum_{\xi\in\mathbb{Z}} \widehat{f}(\xi) e^{-\xi^2 t} e^{i\xi x}$$

where $\widehat{f}(\xi) = \langle f, e^{ix\xi}\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-ix\xi}dx$. Note that if $f \in C^1(T)$ (or even f is continuous and piecewise C^1 on T, meaning f' has only a finite number of jump discontinuities) then $\widehat{f} \in l^1(\mathbb{Z})$; then this series for $u(x,t)$ converges absolutely and uniformly for $x \in T$ and $t \geq 0$, and $u(x,0) = f(x)$; moreover, for $t > 0$, this is a C^∞ solution of $u_t = u_{xx}$. This is because $e^{-\xi^2 t}$ decays very rapidly as $|\xi| \to \infty$ for $t > 0$. But for $t < 0$, we do not expect this series to converge unless $|\widehat{f}(\xi)| \to 0$ extremely fast as $|\xi| \to \infty$. These properties are common for *parabolic equations*: the solution is smooth for $t > 0$, but we *cannot* go backwards in time.

Remark. As for the wave equation, we can also solve IBVP of the form

$$\begin{array}{llll} \text{DE} & u_t = u_{xx} & (0 \leq x \leq \pi, t \geq 0) & \\ \text{IC} & u(x,0) = f(x) & (0 \leq x \leq \pi) & \\ \text{BC} & u(0,t) = 0, & u(\pi,t) = 0 & (t \geq 0) \end{array}$$

(or with BC $u_x(0,t) = 0$, $u_x(\pi,t) = 0$, etc.)

Final Comment. The partial sums $S_k(f) = \sum_{\xi=-k}^{k} \widehat{f}(\xi)e^{ix\xi}$ of the Fourier series of f are obtained by convolving f with the "Dirichlet kernel" $D_k(x) = \sum_{\xi=-k}^{k} e^{ix\xi} = \frac{\sin\left(\left(k+\frac{1}{2}\right)x\right)}{\sin\left(\frac{1}{2}x\right)}$: $\widehat{f * D_k} = \widehat{f}\widehat{D}_k$, so $f * D_k(x) = \sum_{\xi\in\mathbb{Z}} \widehat{f}(\xi)\widehat{D}_k(\xi)e^{ix\xi} = \sum_{\xi=-k}^{k} \widehat{f}(\xi)e^{ix\xi} = S_k(f)$. The Dirichlet kernel, however, is *not* a summability kernel: D_k is not nonnegative (not horrible), and it does not satisfy condition (4) of a summability kernel.

Our next main topic is Fourier Transforms. before discussing Fourier Transforms, we will briefly discuss convolutions.

Convolutions

Let f, g be [complex-valued] measurable functions defined on $\mathbb{R}^n$. In general, we define $(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy$ $(x \in \mathbb{R}^n)$ whenever the integral makes sense.

Example. If $f \in L^1(\mathbb{R}^n)$ and $g \in L^\infty(\mathbb{R}^n)$, then $\forall\, x \in \mathbb{R}^n$, $f(x-y)g(y) = L^1(\mathbb{R}^n_y)$, so $(f * g)(x)$ is defined $(\forall\, x \in \mathbb{R}^n)$, and $\|f * g\|_\infty \leq \|f\|_1 \cdot \|g\|_\infty$. Moreover, $f * g$ is continuous since $|(f*g)(x) - (f*g)(z)| \leq \|g\|_\infty \int_{\mathbb{R}^n} |f(x-y) - f(z-y)|dy \to 0$ as $|x - z| \to 0$ by the continuity of translation on L^1. Thus $f * g \in C_b(\mathbb{R}^n)$ (bounded continuous functions on $\mathbb{R}^n$). In fact, $f * g$ is uniformly continuous.

Theorem. (L^1 convolution on $\mathbb{R}^n$). If $f, g \in L^1(\mathbb{R}^n)$, then for a.e. $x \in \mathbb{R}^n$, $f(x-y)g(y) \in L^1(\mathbb{R}^n_y)$, so $f * g$ is defined a.e. Moreover, $f * g \in L^1(\mathbb{R}^n)$, and $\|f * g\|_1 \leq \|f\|_1 \cdot \|g\|_1$.

Proof. By Tonelli, $\int_{\mathbb{R}^n} |f * g(x)|dx \leq \int\int |f(x-y)g(y)|dydx = \int |g(y)| \int |f(x-y)|dxdy = \|g\|_1 \cdot \|f\|_1 < \infty$, so $f(x-y)g(y) \in L^1(\mathbb{R}^n_x \times \mathbb{R}^n_y)$. The rest follows. $\square$

Properties of Convolutions

Commutativity $(f * g)(x) = (g * f)(x)$

(by change of variables $z = x - y$, $\int f(x-y)g(y)dy = \int f(z)g(x-z)dz$)

Associativity $(f * g) * h = f * (g * h)$.

Young's Inequality. Suppose $1 \leq p, q, r \leq \infty$ and $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$. If $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, then $f * g$ is defined a.e., $f * g \in L^r(\mathbb{R}^n)$, and $\|f * g\|_r \leq \|f\|_p \cdot \|g\|_q$.

Proof. The case $r = \infty$ follows from Hölder's Inequality (since then $\frac{1}{p} + \frac{1}{q} = 1$): $|f * g(x)| \leq \int |f(x-y)g(y)|dy \leq \|f\|_p \cdot \|g\|_q$. Moreover, in this case $f*g$ is uniformly continuous (one of p, q is $< \infty$; by commutativity WLOG $p < \infty$; $|f*g(x) - f*g(z)| \leq \|g\|_q \|f(x-\cdot) - f(z-\cdot)\|_p \to 0$ as $|x - z| \to 0$ by continuity of translation on L^p (as $p < \infty$), so $f * g \in C_b(\mathbb{R}^n)$.

When $r < \infty$, then also $p, q < \infty$. If either p or q is 1 (say WLOG $q = 1$), then Minkowski's Inequality for Integrals ($\left\|\int h(\cdot, y)dy\right\|_p \leq \int \|h(\cdot, y)\|_p dy$, see Jones §11E) implies $\|f * g\|_p \leq \left\|\int |f(\cdot - y)g(y)|dy\right\|_p \leq \int \|f(\cdot - y)g(y)\|_p dy = \|f\|_p \|g\|_1$. The last case is $r < \infty$ and $1 < p, q < \infty$ (which also implies $r > 1$). Let p', q' be the exponents conjugate to p, q, respectively. Then $\frac{1}{p'} = 1 - \frac{1}{p}$, $\frac{1}{q'} = 1 - \frac{1}{q}$. It follows that $\frac{1}{r} + \frac{1}{p'} + \frac{1}{q'} = 1$, and quick calculations give $\left(1 - \frac{p}{r}\right) q' = p$, $\left(1 - \frac{q}{r}\right) p' = q$. By Hölder's Inequality for three functions (see problem 2 in Jones §10A),

$$\begin{aligned}
|f * g(x)| &\leq \int (|f(x-y)|^{\frac{p}{r}} |g(y)|^{\frac{q}{r}}) |f(x-y)|^{\left(1-\frac{p}{r}\right)} |g(y)|^{\left(1-\frac{q}{r}\right)} dy \\
&\leq \left(\int |f(x-y)|^p |g(y)|^q dy\right)^{\frac{1}{r}} \left(\int |f(x-y)|^{\left(1-\frac{p}{r}\right)q'} dy\right)^{\frac{1}{q'}} \left(\int |g(y)|^{\left(1-\frac{q}{r}\right)p'} dy\right)^{\frac{1}{p'}} \\
&= [(|f|^p * |g|^q)(x)]^{\frac{1}{r}} \|f\|_p^{\frac{p}{q'}} \|g\|_q^{\frac{q}{p'}}.
\end{aligned}$$

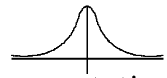
So $\int |f * g(x)|^r dx \leq \| \, |f|^p * |g|^q \|_1 \|f\|_p^{\frac{rp}{q'}} \|g\|_q^{\frac{rq}{p'}} \leq \| \, |f|^p \, \|_1 \, \| \, |q|^q \, \|_1 \, \|f\|_p^{rp/q'} \, \|g\|_q^{rq/p'} = \|f\|_p^{p+rp/q'} \|g\|_q^{q+rq/p'}$ and the result follows since $p + \frac{rp}{q'} = q + \frac{rq}{p'} = r$. $\square$

Note the special cases:

(i) if $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leq p, q \leq \infty$, $\|f * g\|_\infty \leq \|f\|_p \|g\|_q$

(ii) if $1 \leq p \leq \infty$, $\|f * g\|_p \leq \|f\|_p \|g\|_1$

Approximate [Convolution] Identities (or convolution with approximate δ-functions).

Suppose $g \in L^1(\mathbb{R}^n)$ is "peaked" near the origin (like  $\mathbb{R}_n$) and $\int_{\mathbb{R}^n} g = 1$. Then we expect $(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy \approx f(x)$. This expectation can be made rigorous by letting g become more "peaked" as follows: choose any $\varphi \in L^1(\mathbb{R}^n)$ for which $\int_{\mathbb{R}^n} \varphi(x)dx = 1$; for $\varepsilon > 0$, let $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right)$; then by change of variables, also $\int_{\mathbb{R}^n} \varphi_\varepsilon(x)dx = 1$ for all $\varepsilon > 0$. For any fixed $\delta > 0$, $\int_{|x|\le\delta} \varphi_\varepsilon(x)dx = \int_{|x|\le\frac{\delta}{\varepsilon}} \varphi(x)dx$, so $\int_{|x|\le\delta} \varphi_\varepsilon(x)dx \to 1$ as $\varepsilon \to 0$; in this sense φ_ε gets more "peaked" as $\varepsilon \to 0$. The family of functions $\{\varphi_\varepsilon : \varepsilon > 0\}$ (or $\{\varphi_\varepsilon : 0 < \varepsilon \le \varepsilon_0\}$ for some $\varepsilon_0 > 0$, or $\{\varphi_{\varepsilon_j} : j = 1, 2, \ldots\}$ for some sequence $\varepsilon_j \to 0$) is called an *approximate δ-function* or an *approximate identity* (for convolution). The latter name is clarified by the following theorem. To allow the case of a continuum of values of ε, we first need:

Extension of the Lebesgue Dominated Convergence Theorem

Let A be a metric space and α_0 be a limit point of A. Suppose E is a measurable subset of $\mathbb{R}^n$, and $g \in L^1(E)$. Suppose

(i) for each $\alpha \ne \alpha_0 \in A$, $f_\alpha : E \to \mathbb{C}$ is measurable, and $|f_\alpha(x)| \le g(x)$ a.e.

(ii) for some measurable function $f : E \to C$, $\lim_{\alpha\to\alpha_0} f_\alpha(x) = f(x)$ a.e.

Then $f \in L^1(E)$, and $\lim_{\alpha\to\alpha_0} \int_E f_\alpha(x)dx = \int_E f(x)dx$.

Proof. If $\int_E f_\alpha \not\to \int_E f$, then $\exists$ a sequence $\alpha_n \to \alpha_0$ and an $\varepsilon > 0$ for which $(\forall\, n)\ \left|\int_E f_{\alpha_n} - \int_E f\right| \ge \varepsilon$. But by the LDCT, $\int_E f_{\alpha_n} \to \int_E f$. Contradiction. $\square$

Theorem. Suppose $\varphi \in L^1(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} \varphi = 1$. For $\varepsilon > 0$, let $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right)$.

(a) If $f \in L^p(\mathbb{R}^n)$ where $1 \le p < \infty$, then $f * \varphi_\varepsilon \to f$ in L^p (i.e., $\|f * \varphi_\varepsilon - f\|_p \to 0$).

(b) $(p = \infty)$. If f is bounded and uniformly continuous on $\mathbb{R}^n$, then $f * \varphi_\varepsilon \to f$ uniformly.

Proof.

(a) Let $\psi(y) = \|f_y - f\|_p$ for $y \in \mathbb{R}^n$ where $f_y(x) = f(x-y)$. Then ψ is continuous, ≥ 0, bounded by $2\|f\|_p$, and $\psi(0) = 0$. By Minkowski's Inequality for Integrals, for $\varepsilon > 0$,

$$\begin{aligned} \|f * \varphi_\varepsilon - f\|_p &= \left\| \int_{\mathbb{R}^n} (f(\cdot - y) - f(\cdot))\varphi_\varepsilon(y)dy \right\|_p \\ &\le \int_{\mathbb{R}^n} \|(f(\cdot - y) - f(\cdot))\varphi_\varepsilon(y)\|_p dy \\ &= \int_{\mathbb{R}^n} \psi(y)|\varphi_\varepsilon(y)|dy = \int_{\mathbb{R}^n} \psi(\varepsilon z)|\varphi(z)|dz \to 0 \end{aligned}$$

as $\varepsilon \to 0$ by the extended LDCT.

(b) The proof in part (a) fails if only $f \in L^\infty$, but goes through with $p = \infty$ when f is bounded and unif. cont.

$\square$

Mollification. One of the main applications of this theorem is when φ is smooth. As we will see, mild further assumptions on φ and f imply that $f * \varphi_\varepsilon$ is as smooth as φ. So $f * \varphi_\varepsilon$ is a smooth function close to f in L^p.

Notation: A multi-index is an $\alpha \in \mathbb{Z}^n$ with each $\alpha_i \geq 0$. For $x \in \mathbb{R}^n$, $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{a_n}$. Define $\partial_x^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\partial_n}$. We write $|\alpha| = \sum_{i=1}^n \alpha_i$.

Differentiation Under the Integral Sign. Suppose E is a measurable subset of $\mathbb{R}^n$, and $I \subset \mathbb{R}$ is an interval. Suppose $f : E \times I \to \mathbb{C}$ satisfies $(\forall\, t \in I) f(\cdot, t) \in L^1(E)$, $\frac{\partial f}{\partial t}(x,t)$ exists for all $(x,t) \in E \times I$, and $\exists\, g \in L^1(E) \ni (\forall (x,t) \in E \times I) \left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x)$. Let $F(t) = \int_E f(x,t)dx$. Then F is differentiable on I, and $F'(t) = \int_E \frac{\partial f}{\partial t}(x,t)dx$.

Proof. Fix $t_0 \in I$. For $t \neq t_0 \in I$, let $h(x,t) = \frac{f(x,t) - f(x,t_0)}{t - t_0}$. The Mean Value Theorem implies $|h(x,t)| \leq \sup_{t \in I} \left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x)$, so the extended LDCT implies $\lim_{t \to t_0} \frac{F(t) - F(t_0)}{t - t_0} = \lim_{t \to t_0} \int_E h(x,t)dx = \int_E \lim_{t \to t_0} h(x,t)dx = \int_E \frac{\partial f}{\partial t}(x,t)dx$. $\square$

Remark. A similar result is true for partial derivatives when $I^{\text{open}} \subset \mathbb{R}^m$.

Differentiating Convolutions

Theorem. Suppose $f \in L^1(\mathbb{R}^n)$, $g \in C^k(\mathbb{R}^n)$, and for $|\alpha| \leq k$, $\partial^\alpha g$ is bounded. Then $f * g \in C^k(\mathbb{R}^n)$, and for $|\alpha| \leq k$, $\partial^\alpha (f * g) = f * \partial^\alpha g$.

Proof. Use induction on $|\alpha|$. For $|\alpha| = 1$, write $(f * g)(x) \int_{\mathbb{R}^n} g(x-y) f(y) dy$. Then $\partial_x^\alpha (g(x-y)f(y)) = \partial^\alpha g(x-y) f(y)$, so $\|\partial^\alpha g\|_\infty |f(y)|$ is a dominating function, and we can differentiate under the integral sign. $\square$

Corollary. If $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and $g \in C_c^k(\mathbb{R}^n)$ (c for compact support), then $f * g \in C^k(\mathbb{R}^n)$.

Proof. For $R > 0$, let $B_R = \{x : |x| < R\}$. Suppose $g(x) \equiv 0$ for $|x| \geq R$. For $|x| < N$, $f * g(x) = \int_{B_R} f(x-y) g(y) dy = \int_{B_R} (\psi f)(x-y) g(y) dy = (\psi f) * g(x)$, where $\psi = B_{N+R}$. Since $\psi f \in L^1$, the corollary follows. $\square$

Note. If $f \in L^p_{\text{loc}}(\mathbb{R}^n)$ for any p $(1 \leq p \leq \infty)$, then $f \in L^1_{\text{loc}}(\mathbb{R}^n)$.

Remark on supports: Clearly $\operatorname{supp}(f * g) \subset \operatorname{supp}(f) + \operatorname{supp}(g)$.

Theorem. Let Ω be an open subset of $\mathbb{R}^n$ and $1 \leq p < \infty$. Then $C_c^\infty(\Omega)$ (C^∞ functions with compact support in Ω) is dense in $L^p(\Omega)$.

Proof. Let $K_1 \subset K_2 \subset \cdots$ be a compact exhaustion of Ω, i.e., each K_j is a compact subset of Ω, $K_j \subset K_{j+1}^0$ (interior), and $\bigcup K_j = \Omega$. Given $f \in L^p(\Omega)$ and $\varepsilon > 0$, the LDCT implies

$f\chi_{K_j} \to f$ in $L^p(\Omega)$, so we can choose j for which $\|f\chi_{K_j} - f\|_p < \frac{\varepsilon}{2}$. Since $K_j \subset K^0_{j+1}$, $\exists\, \eta > 0 \ni K_j + B_\eta \subset K^0_{j+1}$. Let $\varphi(x)$ be in $C_c^\infty(\mathbb{R}^n)$ with $\varphi(x) = 0$ for $|x| \geq 1$ and $\int_{\mathbb{R}^n} \varphi(x)dx = 1$ (e.g., $\varphi(x) = \frac{\psi(x)}{\int \psi}$ where $\psi(x) = \exp\left(\frac{1}{1-|x|^2}\right)$ for $|x| < 1$ and $\psi(x) = 0$ for $|x| \geq 1$), and let $\varphi_\delta(x) = \delta^{-n}\varphi\left(\frac{x}{\delta}\right)$ for $0 < \delta \leq \eta$. Now $f\chi_{K_j} * \varphi_\delta \to f\chi_{K_j}$ in $L^p(\mathbb{R}^n)$ as $\delta \to 0$. Since $\operatorname{supp}(f\chi_{K_j} * \varphi_\delta) \subset \Omega$ for $0 < \delta \leq \eta$, $\exists\, \delta \in (0, \eta]$ for which $\|f\chi_{K_j} * \varphi_\delta - f\chi_{K_j}\|_p < \frac{\varepsilon}{2}$, and $f\chi_{K_j} * \varphi_\delta \in C_c^\infty(\Omega)$. $\square$