Article

The Momentum Operator in The Infinite Square Well Problem of Quantum Mechanics

Daniel Abramson*

9/72 Soi Kasemsan 3, Rama 1 Road, Wangmai, Pathumwan, Bangkok 10330

*Corresponding author E-mail: khunchang@gmail.com

Received: November 20th, 2020, Revised: December 25th, 2020, Accepted: December 26th, 2020

Abstract

The infinite square well (ISW) is usually the first problem involving bound states studied in quantum mechanics. Although it is straightforward to determine the stationary states and energy levels, the artificial structure of the problem leads to simple paradoxes involving quantization and measurability of momentum. We show how the boundary conditions influence the momentum operator P, and discover the connections between momentum, differentiation and integration, which rescue the hermicity of P and its powers P^n.

Keywords: Infinite Square Well, Momentum Operator

The one-dimensional infinite square well (ISW) is discussed in virtually every quantum mechanics text as the first example of how quantization arises from boundary conditions. Yet the ISW problem gives rise to conceptual paradoxes which can be a distraction at the initial stage of development of the theory. For example, while energy in the ISW is quantized, momentum is neither quantized nor measurable (even in principle), shattering the classical relation $E = p^2/2m$. After explaining these paradoxes, we deepen our understanding of the role of the boundary conditions by exploring the relationship between hermicity of P and the integration by parts formula. While hermicity of P and P^2 are easy to establish owing to fortunate happenstances, when we try to show that P^3 and higher powers are Hermitian, we are forced to amend the conventional expressions for derivatives of the wave function.

A Paradox

The one-dimensional infinite square well problem is defined by the potential

$$V(x) = \begin{cases} 0, & 0 < x < L \\ \infty, & \text{otherwise} \end{cases}.$$

The energy levels for a particle with mass m confined in an ISW of width L are quantized [1] with

$$E_n = \frac{\hbar^2\pi^2 n^2}{2mL^2}.$$

If we observe and measure the particle's energy, we obtain one of the eigenvalues E_n of the Hamiltonian operator $H = P^2/2m + V(X)$, and the state vector collapses to the energy eigenstate $|E_n\rangle$.[a]

A particle in an energy eigenstate $|E_n\rangle$ would seem to have a definite, quantized momentum $p_n = \pm\sqrt{2mE_n} = \pm\hbar\pi n/L$. Therefore, $|E_n\rangle$ should also be a momentum eigenstate $|p_n\rangle$ with eigenvalues $p_n = \pm\hbar\pi n/L$. If $|p_n\rangle$ is a momentum eigenstate, then $(\Delta P)^2 = \langle P^2\rangle - \langle P\rangle^2 = \langle p_n|P^2|p_n\rangle - \langle p_n|\langle P\rangle^2|p_n\rangle = p_n^2\langle p_n|p_n\rangle - p_n^2\langle p_n|p_n\rangle = 0$, which would be consistent with the wave function (in p-space) of a momentum eigenstate, $\langle p|p_n\rangle = \delta(p - p_n)$. By the uncertainty principle, therefore, $\Delta X = \infty$. But $\Delta X \leq L$ since the particle is confined in the well — a contradiction. This paradox is based on several misconceptions.

The Momentum Operator $P = -i\hbar D_x$

Energy eigenstates are not momentum eigenstates. To see this, consider a particle confined within $0 \leq x \leq L$. The energy eigenfunctions $E_n(x) \equiv \langle x|E_n\rangle$ in the x-basis [1] are:

$$E_n(x) = \begin{cases} \sqrt{2/L}\sin\frac{n\pi x}{L}, & 0 \leq x \leq L \\ 0, & \text{otherwise} \end{cases}. \tag{1}$$

Therefore, on the open interval $(0, L)$:

$$\langle x|P|E_n\rangle = -i\hbar\frac{d}{dx}\left(\sqrt{2/L}\sin\frac{n\pi x}{L}\right) = \frac{\hbar n\pi}{iL}\sqrt{2/L}\cos\frac{n\pi x}{L} \neq \lambda\sqrt{2/L}\sin\frac{n\pi x}{L} = \langle x|\lambda|E_n\rangle.$$

Since $P|E_n\rangle \neq \lambda|E_n\rangle$ for any number λ, energy eigenstates are not momentum eigenstates.

---

[a] It is not necessary to make an arbitrarily precise measurement. If the precision of the measurement is better than $(E_n - E_{n-1})/2$, we know the particle's energy is exactly E_n and its state vector is $|E_n\rangle$ immediately after the measurement is made.

Momentum eigenstates $|p\rangle$ do not exist. Such a state would violate the uncertainty principle, as explained above. Alternatively, suppose P has an eigenvector $|p\rangle$; say $P|p\rangle = p|p\rangle$ where p is a real number.[b] The wave function $\psi_p(x) \equiv \langle x|p\rangle$ in the x-basis satisfies:

$$-i\hbar\psi_p'(x) = p\psi_p(x)\,. \tag{2}$$

The solution to (2) is $\psi_p(x) = (2\pi\hbar)^{-1/2}e^{ipx/\hbar}$, normalized to $\langle p'|p\rangle = \delta(p - p')$. But $\psi_p(x)$ is not supported[c] in $[0, L]$; moreover, $\psi_p(x)$ is never zero, so it cannot satisfy the boundary conditions. Therefore, $|p\rangle$ is not a possible ISW state vector.

More generally, no state $|\psi\rangle$ is allowed whose wave function in p-space $\hat{\psi}(p) \equiv \langle p|\psi\rangle$ is too "sharply peaked." Specifically, $\Delta X \le L$ implies $\Delta P \ge \hbar/2\Delta X \ge \hbar/2L$. This means ISW state vectors must have well-spread wave functions in p-space. The narrower the well, the larger the variance in $\hat{\psi}(p)$.

Momentum cannot be measured, even in principle. If momentum could be measured with a high degree of precision, the state vector would collapse to an eigenvector $|p\rangle$ or to a state with a sharply-peaked wave function in p-space, which we have seen is not possible. Nevertheless, any state can be expressed as a superposition of momentum pseudoeigenstates $|\psi\rangle = \int\langle p|\psi\rangle|p\rangle\, dp$ — "pseudo" because the individual eigenstates $|p\rangle$ are not found in the ISW. In x-space, this superposition is given by

$$\psi(x) = \langle x|\psi\rangle = \int_{-\infty}^{\infty}\langle x|p\rangle\langle p|\psi\rangle\, dp = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} e^{ipx/\hbar}\hat{\psi}(p)\, dp \tag{3}$$

where

$$\hat{\psi}(p) = \langle p|\psi\rangle = \int_{-\infty}^{\infty}\langle p|x\rangle\langle x|\psi\rangle\, dx = \int_0^L\langle x|p\rangle^*\psi(x)\, dx = \frac{1}{\sqrt{2\pi\hbar}}\int_0^L e^{-ipx/\hbar}\psi(x)\, dx\,. \tag{4}$$

These equations recapitulate the relationship between $\langle x|\psi\rangle$ and $\langle p|\psi\rangle$ as Fourier transforms.

Unique energy without unique momentum. A particle in an energy eigenstate $|E_n\rangle$ has a unique energy E_n, but is a superposition of states with different momenta. We conclude that $E = p^2/2m$ is not a correct or meaningful equation for the particle. While the pseudoeigenstates $|p\rangle$ are not found in the ISW, the particle has a momentum probability density $|\langle p|\psi\rangle|^2$ given by (4).

[b] Could P have an imaginary eigenvalue? Suppose $P|p\rangle = ip|p\rangle$ for some $p \in \mathbb{R}$. Then the solution to (2) is $\psi_p(x) = Ae^{-px/\hbar}$. Such a wave function cannot be normalized on $(-\infty, \infty)$ either to $\langle p'|p\rangle = 1$ or $\langle p'|p\rangle = \delta(p - p')$, so we rule this out.

[c] The *support* of a function f, written $\mathrm{supp}(f)$, is the set of x for which $f(x) \neq 0$.

How can we understand the statement "momentum cannot be measured"? Momentum, unlike energy and position, cannot be measured with arbitrarily high precision. Can we understand, physically, why momentum is resistant to measurement? Yes: one way to measure a particle's momentum with arbitrarily high precision is to scatter off the particle a photon with arbitrarily low frequency. (Shankar [2] (p. 123) shows how to reconstruct the initial and final momenta p and p' of a particle after collision with a photon, as a function of the incoming and outgoing frequencies of the photon ω and ω', both of which can be known exactly in principle. One finds that $\omega \to 0$ implies $\omega' \to 0$ for any choice of p, which in turn implies that $p - p'$ can be made as small as desired.) But the frequency of a photon trapped inside an ISW of width L is quantized with $\omega = n\pi c/L$ (see Garrison and Chiao [3], Eq. 2.15). Therefore, the photon's maximum wavelength is $\lambda = 2\pi c/\omega_{\min} = 2L$; so we cannot make a measurement in the ISW with photons of arbitrarily long wavelength. (Fulling[d] [4] discusses the momentum measurement problem in a wider experimental context.)

Energy eigenfunctions in momentum space. To penetrate the conundrum of an energy eigenstate being a superposition of momentum states, it is instructive briefly to examine the momentum probability density $|\langle p|E_n\rangle|^2$. (Liang, Zhang and Dardenne [5] provide a thorough summary of momentum distributions in the ISW problem.) A particle in an energy eigenstate $|E_n\rangle$ has a wave function in p-space:

$$\hat{E}_n(p) = \langle p|E_n\rangle = \int_{-\infty}^{\infty} \langle p|x\rangle\langle x|E_n\rangle\, dx = \frac{1}{\sqrt{\pi\hbar L}} \int_0^L e^{-ipx/\hbar} \sin\frac{n\pi x}{L}\, dx\, .$$

Evaluation of the integral gives

$$\hat{E}_n(p) = \sqrt{2\alpha/\pi}\, i^{n+1} e^{-ip\alpha} \left\{(-1)^n \frac{\sin\alpha(p+p_n)}{2\alpha(p+p_n)} - \frac{\sin\alpha(p-p_n)}{2\alpha(p-p_n)}\right\}$$

where $\alpha \equiv L/2\hbar$ and $p_n \equiv \sqrt{2mE_n} = n\pi\hbar\big/L = n\pi\big/2\alpha$. Squaring gives:

$$\left|\hat{E}_n(p)\right|^2 = \frac{1}{2\pi\alpha}\left\{\frac{\sin^2[\alpha(p+p_n)]}{(p+p_n)^2} + \frac{\sin^2[\alpha(p-p_n)]}{(p-p_n)^2} - (-1)^n \frac{2\sin[\alpha(p-p_n)]\sin[\alpha(p+p_n)]}{(p-p_n)(p+p_n)}\right\}.$$

This function has two peaks at $p = \pm p_n$ (except for $n = 1$). As $n \to \infty$:

---

[d] Fulling addresses the non-measurability of momentum in the ISW by *gedanken* experiments such as measuring time-of-flight, slit-and-screen diffraction and bending of paths in a magnetic field, as well as scattering with other particles, and notes that "the localized nature of the experiments makes precise momentum measurements impossible."

a) If α is held fixed, the shape of the peaks at $\pm p_n = \pm n\pi/2\alpha$ remains almost the same, but the peaks move further apart with increasing n (due to increasing energy).

b) If $n/\alpha = (2/\pi)p_n$ is held fixed, the peaks grow higher and narrower. This makes sense since $\alpha \to \infty$ $(L \to \infty)$ approaches the case of the free particle.

Figures 1 and 2 show $\left|\hat{E}_n(p)\right|^2$ for $n = 10$, and $\alpha = 10$ and $\alpha = 20$.

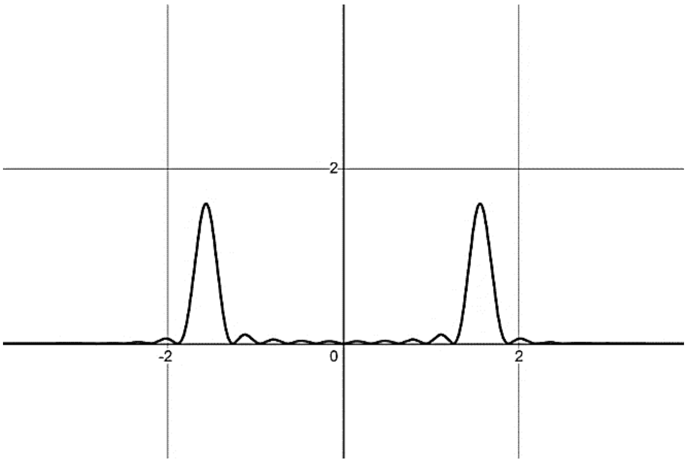

Figure 1 *The momentum probability density* $\left|\hat{E}_n(p)\right|^2$ *of a particle in an energy eigenstate* $|E_n\rangle$ *for* $n = 10$ *and* $\alpha = 10$

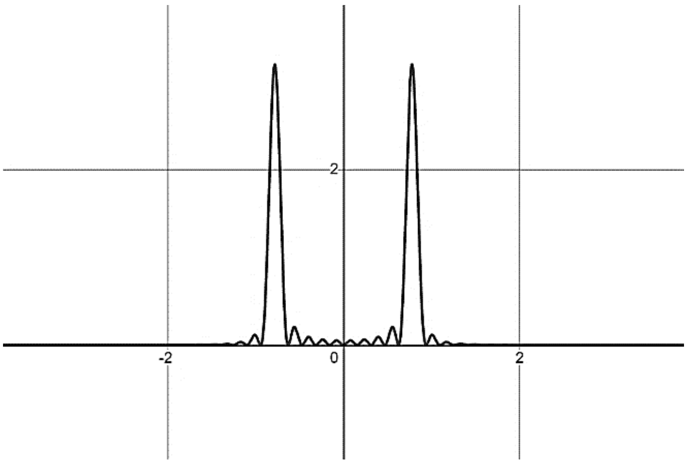

Figure 2 *The momentum probability density* $\left|\hat{E}_n(p)\right|^2$ *of a particle in an energy eigenstate* $|E_n\rangle$ *for* $n = 10$ *and* $\alpha = 20$

Hermicity of P

It is a postulate of quantum mechanics that observable quantities are representable as Hermitian operators [1][2]. We will see that hermicity of P is equivalent to integration by parts (with vanishing of the boundary term), which in turn relies upon a function being the integral of its derivative.

The fundamental theorem of calculus $\int_a^x f'(t)\,dt = f(x) - f(a)$ is usually proved when f is differentiable and f' is continuous everywhere in the domain of f. However, the energy eigenfunctions $E_n(x)$ in (1) are not differentiable at $x = 0$ and $x = L$: $E_n(x)$ has a cusp at the boundaries so $E_n'(x)$ has jump discontinuities at these points. Indeed, if $|\psi\rangle$ is any state vector in the ISW with wave function $\psi(x) = \langle x|\psi\rangle$, its derivative $\psi'(x)$ must be discontinuous at $x = 0$ and $x = L$.[e] We want to apply the fundamental theorem of calculus to discontinuous functions like these; therefore, we need to know the weakest conditions under which a function is the integral of its derivative.

The theory of Lebesgue integration introduces the property of *absolute continuity* of a function, which is stronger than pointwise or uniform continuity.[f] It is a theorem (Rudin [6], Theorem 7.18) that $f\colon [a,b] \to \mathbb{C}$ is absolutely continuous if and only if f has a derivative f' "almost everywhere", f' is integrable and

$$f(x) = f(a) + \int_a^x f' \tag{5}$$

for all $a \le x \le b$. In other words, f is the integral of its derivative. ("Almost everywhere" means everywhere except on a set of measure zero. A set E has *measure zero* if it can be covered by a finite or countable union of intervals (α_i, β_i) whose total length $\sum(\beta_i - \alpha_i)$ can be made arbitrarily small. In other words, for any $\epsilon > 0$ there exists a countable collection of intervals $I_i = (\alpha_i, \beta_i)$ such that $E \subseteq \bigcup I_i$ and $\sum(\beta_i - \alpha_i) < \epsilon$.)

Lebesgue measure and the Lebesgue integral are building blocks of integration theory (Rudin [6] is the standard reference). We will not require any deep results of this theory here, apart from basic terminology and the theorem stated in (5), which we will take as the definition of absolute continuity. We will assume that all wave functions and their derivatives are absolutely continuous except at a finite

[e] Note that $\psi'(x) = 0$ for $x \le 0$ and $x \ge L$. If ψ' is continuous at $x = 0$, then $\psi'(0) = 0$. Hence $\psi(0) = \psi'(0) = 0$. These initial conditions force the unique solution to any equation of the form $\psi''(x) + a(x)\psi'(x) + b(x)\psi(x) = c(x)$ to be the trivial $\psi(x) = 0$. The Schrödinger equation $H\psi = E\psi$ has this form. $\psi(x) = 0$ is not a physical solution to the Schrödinger equation since it cannot be normalized. Hence $\psi(0) = \psi'(0)$ is not possible. The same applies at $x = L$.

[f] A function $f\colon [a,b] \to \mathbb{C}$ is *absolutely continuous* on $[a,b]$ if for every $\epsilon > 0$ there exists $\delta > 0$ such that if $(\alpha_1, \beta_1), \dots, (\alpha_n, \beta_n)$ is any pairwise disjoint collection of intervals within $[a,b]$ whose lengths satisfy $\sum_{i=1}^n (\beta_i - \alpha_i) < \delta$, then $\sum_{i=1}^n |f(\beta_i) - f(\alpha_i)| < \epsilon$.

number of points (specifically, two). While there exist continuous functions which are differentiable almost everywhere and do not satisfy (5), these tend to be pathological and would not occur in physical problems.[g] There also exist continuous functions which are not differentiable on a set of positive measure (e.g., an interval); there are even continuous functions which are nowhere differentiable! We will ignore such pathological examples and assume that all solutions to Schrödinger's equation are absolutely continuous, and therefore differentiable almost everywhere and satisfy (5).

Remarks:

a) It is easy to see that a function v with a jump discontinuity cannot satisfy (5). Suppose $v(x)$ has a jump discontinuity at $x = c$. Then

$$\lim_{\epsilon\downarrow 0}\int_{c-\epsilon}^{c+\epsilon} v'(x)\,dx = 0 \neq v(c^+) - v(c^-) = \lim_{\epsilon\downarrow 0}\big(v(c+\epsilon) - v(c-\epsilon)\big)\,.$$

Therefore, for some $\epsilon > 0$:

$$\int_{c-\epsilon}^{c+\epsilon} v' \neq v(c+\epsilon) - v(c-\epsilon)\,.$$

If $v\colon [a,b] \to \mathbb{C}$ is continuous except for a jump discontinuity at $x = c$ $(a \le c \le b)$, then

$$\begin{aligned}\int_a^b v'(x)\,dx &= \left(\int_a^{c-\epsilon} + \int_{c-\epsilon}^{c+\epsilon} + \int_{c+\epsilon}^{b}\right) v'(x)\,dx \\ &\neq \big(v(c-\epsilon) - v(a)\big) + \big(v(c+\epsilon) - v(c-\epsilon)\big) + \big(v(b) - v(c+\epsilon)\big) \\ &= v(b) - v(a)\,.\end{aligned}$$

b) Suppose $u, v : \mathbb{R} \to \mathbb{C}$ and uv is discontinuous at $x = c$. (This requires only that v is discontinuous at $x = c$ and $u(c) \neq 0$.) Then by the previous result, $\int_a^b (uv)' \neq u(b)v(b) - u(a)v(a)$. Therefore, we see that uv is (absolutely) continuous on $[a,b]$ if and only if integration by parts is valid.

---

[g] The Cantor function $\xi(x)$ is a continuous, monotone-increasing function which maps $[0,1]$ to $[0,1]$, is differentiable almost everywhere, and satisfies $\xi'(x) = 0$ (see Rudin [6], Example 7.16(b)). Define:

$$f(x) = \begin{cases} \xi(2x)\,, & \text{if } 0 \le x \le 1/2 \\ \xi(2-2x)\,, & \text{if } 1/2 \le x \le 1 \\ 0\,, & \text{otherwise} \end{cases}$$

Then f is continuous on $\mathbb{R}$, satisfies $f(0) = f(1) = 0$, and $\langle Pf|g\rangle = 0$ for every $|g\rangle$. The Cantor function is uniformly but not absolutely continuous.

Theorem 1: P is Hermitian.

Proof: Let $|f\rangle$and $|g\rangle$ be arbitrary ISW state vectors, and $f(x)$ and $g(x)$ their wave functions in x-space. Note that continuity plus $\mathrm{supp}(f) \subseteq [0, L]$ implies that $f(0) = f(L) = 0$. We would like to argue:

$$\langle f|Pg\rangle = \int_{-\infty}^{\infty} f^*[-i\hbar g'] \stackrel{?}{=} -i\hbar f^* g\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f'^*[-i\hbar g] = \int_{-\infty}^{\infty} [-i\hbar f']^* g = \langle Pf|g\rangle\,. \quad (6)$$

Since f^*g is continuous on $\mathbb{R}$, the integration by parts is valid. Therefore $\langle P^\dagger f\big|g\rangle \equiv \langle f|P|g\rangle = \langle f|Pg\rangle = \langle Pf|g\rangle$, which implies $P^\dagger = P$.

Remarks:

a) There is no contradiction in the fact that momentum p is not measurable (observable) while the operator P is Hermitian. *It is a postulate of quantum mechanics that observable quantities correspond to Hermitian operators, but the converse need not hold.*

b) The relation $\langle f|Pg\rangle = \langle Pf|g\rangle$ holds so long as one of the two functions f, g vanishes at $\pm\infty$, since only one function need vanish to make the boundary term in (6) vanish.

c) Eq. (6) shows that, generally speaking (not just for the ISW problem), hermicity of P is equivalent to the integration by parts formula. Consequently, P is Hermitian on any collection of state vectors $|\psi\rangle$ whose wave functions $\psi(x) = \langle x|\psi\rangle$ are absolutely continuous (guarantees that integration by parts is valid) and satisfy $\psi(\pm\infty) = 0$ (boundary term vanishes). Note that $\int_{-\infty}^{\infty}|\psi(x)|^2\, dx < \infty$ does not imply that $\psi(\pm\infty) = 0$, as the example $\psi(x) = \exp(-x^2 \sin^2 \pi x)$ in Figure 3 demonstrates.

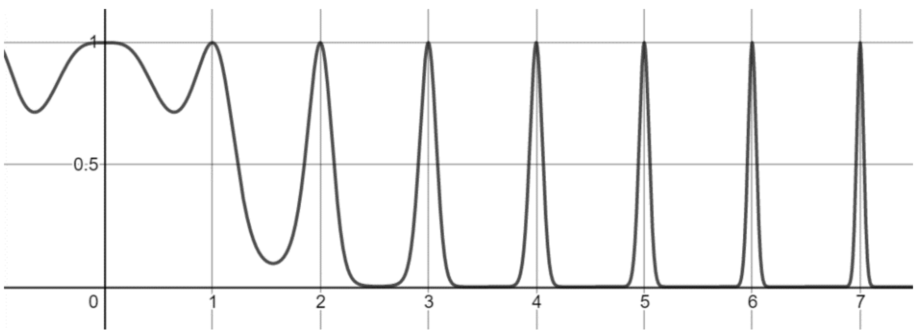

Figure 3 *The function* $\psi(x) = \exp(-x^2 \sin^2 \pi x)$ *is square integrable but does not vanish at* $\pm\infty$

d) Eq. (6) can hold with both sides $= \pm\infty$. Example: let $f(x) = \sqrt{x(L-x)}$ if $0 \le x \le L$, and $f(x) = 0$ otherwise. Note that f is continuous on $[0, L]$ and $f' \in L^1[0, L]$. But $f' \notin L^2[0, L]$, therefore $\langle Pf|Pf\rangle = \infty$. If $g = Pf$, then $\langle Pf|g\rangle = \langle f|Pg\rangle = \infty$.[h] In this case $2m\langle H\rangle = \langle f|P^2|f\rangle = \infty$, so the particle would have infinite energy, which is not physically realistic.

Hermicity of P^2

Hermicity of $H = P^2/2m + V(X)$ is expected, since H corresponds to an observable (the total energy). Indeed, hermicity of P^2 should follow from hermicity of P, since:

$$\langle [P^2]^\dagger\psi| = \langle\psi|P^2 = \langle P^\dagger\psi|P = \langle P\psi|P = \langle P^\dagger P\psi| = \langle P^2\psi| \,.$$

A possible objection to this argument is that $\langle P\psi|$ is not a bra (equivalently: $|P\psi\rangle$ is not a ket) in the ISW, since the wave function $\langle x|P\psi\rangle = -i\hbar\psi'(x)$ is not defined at $x = 0$ and $x = L$. (This is consistent with the fact that the particle reflects off the boundary and has no well-defined momentum at these points.) So $|P\psi\rangle$ has no projection onto the x-basis vectors $|x = 0\rangle$ and $|x = L\rangle$. Therefore, is the expression $\langle P\psi|P$ well defined?

One might dismiss these objections on the grounds that inner products are integrals $\langle\phi|\psi\rangle = \int_0^L \phi^*(x)\psi(x)\,dx$; therefore, the definition (or lack thereof) of the integrands at finite number of points does not matter. But herein lies the rub. While the discontinuities of $\psi'(x)$ at $x = 0$ and $x = L$ do not affect its properties as an integrand, we will see that $\psi''(x)$ includes factors of $\delta(x)$ and $\delta(x - L)$ to preserve the relation $\int \psi'' = \psi'$, and these *do* affect the properties of ψ'' as an integrand.

To illustrate the problem in concrete and dramatic fashion, consider an ISW with walls at $x = 0$ and $x = 1$. Let $|f\rangle$ and $|g\rangle$ be (unnormalized) state vectors with wave functions $f(x) = (1 - x)\arctan x$ and $g(x) = x(1 - x)$ on the interval $0 \le x \le 1$, and vanishing elsewhere. We calculate:

$$\langle f|Pg\rangle = \langle Pf|g\rangle = \frac{2\pi + \ln 2 - 7}{6i\hbar}$$

$$\langle f|P^2 g\rangle = \langle P^2 f|g\rangle = \frac{1 - \ln 2}{\hbar^2}$$

but:

$$0 = \langle f|P^3 g\rangle \ne \langle P^3 f|g\rangle = \frac{\pi/4 - 1}{i\hbar^3}.$$

This seems to indicate $\langle f|P^3 g\rangle \ne \langle P^3 f|g\rangle$. Moreover, if the particle's state is $|g\rangle$, then we can compute the expected values:

$$\langle H\rangle = \langle g|Hg\rangle = \hbar^2/6m$$

$$\langle H^2\rangle = \langle g|H^2 g\rangle = 0$$

[h] We will deliberately abuse notation and refer to functions Pf and $Pf(x)$ when we mean $\langle x|Pf\rangle$.

which yields the absurd result $\Delta H^2 = \langle H^2 \rangle - \langle H \rangle^2 = -\hbar^4/36m^2$. In addition, $\langle H^2 \rangle = \langle g|H^2|g \rangle = \langle Hg|Hg \rangle = \hbar^4/m^2$; hence $\langle Hg|Hg \rangle \neq \langle g|H^2 g \rangle$, which would suggest that H is not Hermitian. (Belloni and Robinett[i] [7] and Bonneau et al.[j] [8] discuss similar contradictions.) These results indicate that something is wrong with our inner product calculations. In the next section we will identify the oversight when we prove the hermicity of P^n. But first we dispense with the case $n = 2$, which can be done quickly owing to a stroke of luck.

Theorem 2: $P^2 = -\hbar^2 (D_x)^2$ is Hermitian.

Proof: As in (6), we would like to argue that

$$-\frac{1}{\hbar^2}\langle f|P^2 g\rangle = \int_{-\infty}^{\infty} f^* g'' \stackrel{?}{=} f^* g'\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f'^* g' \stackrel{?}{=} -f'^* g\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} f''^* g = -\frac{1}{\hbar^2}\langle P^2 f|g\rangle \,. \qquad (7)$$

Both sets of boundary terms vanish because f^* and g are supported in $[0, L]$. But the integration by parts formula $\int (uv)' = \int uv' + \int u'v$ requires that uv be absolutely continuous. In the first integration by parts $uv = f^* g'$, and in the second integration by parts $uv = f'^* g$. Since f' and g' are discontinuous at $x = 0$ and $x = L$, we are rightly concerned that $f^* g'$ and $f'^* g$ might be discontinuous at the boundaries. Fortunately, f^* and g both vanish at the boundaries; therefore, $f^* g'$ and $f'^* g$ are continuous. Consequently, $f^* g'$ and $f'^* g$ are the integrals of their derivatives, and the integration by parts is correct.

Higher powers of P

Hermicity of P^n should follow from hermicity of P using the relation $(A_1 A_2 \cdots A_n)^\dagger = {A_n}^\dagger \cdots {A_2}^\dagger {A_1}^\dagger$, which implies $P^{n\dagger} = P^{\dagger n} = P^n$. But as noted in the previous section, $|P\psi\rangle$ is not an ISW state vector and the validity of this argument is not clear. In proving that P^2 is Hermitian we were lucky that the boundary terms vanished owing to the presence of at least one factor that vanished at the boundaries of the well. When we try to show that higher powers of P are Hermitian, we find ourselves integrating $\int (uv)' = \int uv' + \int u'v$ where *both factors* in the product uv are discontinuous at the boundaries.

[i] Belloni and Robinett demonstrate the same contradiction when computing the expected value of even powers of the momentum operator $\langle E_n|P^{2q}|E_n\rangle$ in an energy eigenstate, due to incorrect treatment of the boundaries.

[j] Bonneau, Faraut, and Valent distinguish between Hermitian and self-adjoint operators. They state that $P = -i\hbar D_x$ is Hermitian on the set $\mathcal{D}(P) = \{\phi, \phi' \in L^2[0, L];\ \phi(0) = \phi(L) = 0\}$ but not self-adjoint, since $P^\dagger = -i\hbar D_x$ has the same formal expression but acts on a different space of functions (without boundary conditions).

For example, if we try to argue that $\langle P^3 f|g\rangle = \langle P^2 f|Pg\rangle = \langle Pf|P^2 g\rangle = \langle f|P^3 g\rangle$, the middle equality is equivalent to $\int_{-\infty}^{\infty} f''^* g' = -\int_{-\infty}^{\infty} f'^* g''$, which would follow from integrating by parts

$$\int_{-\infty}^{\infty} f''^* g' + \int_{-\infty}^{\infty} f'^* g'' = \int_{-\infty}^{\infty} (f'^* g')' \stackrel{?}{=} f'^* g' \Big|_{-\infty}^{\infty} = 0 \tag{8}$$

if the second equality in (8) were valid. But $f'^* g'$ is generally not continuous at $x = 0$ and $x = L$ since both f'^* and g' are discontinuous at these points. Therefore, (8) is generally false when f'^* and g' represent conventional derivatives.

Remarks:

a) Suppose $\phi: \mathbb{R} \to \mathbb{C}$ is continuous except at $x = c$ where ϕ has a jump discontinuity. We define the "completed derivative":

$$\overline{\phi'}(x) = \phi'(x) + \big(\phi(c^+) - \phi(c^-)\big)\delta(x - c)\,. \tag{9a}$$

This ensures that

$$\lim_{\epsilon \downarrow 0} \int_{c-\epsilon}^{c+\epsilon} \overline{\phi'}(x)\, dx = \phi(c^+) - \phi(c^-)\,.$$

Note that $\overline{\phi'}(x) = \phi'(x)$ everywhere except at $x = c$. The supplementary δ-function "defines" $\phi'(x)$ at $x = c$ so that ϕ is the integral of its derivative. Similarly, if $\phi: \mathbb{R} \to \mathbb{C}$ is supported in $[0, L]$, is continuously differentiable on the open interval $[0, L]$, and has jump discontinuities at $x = 0$ and $x = L$, then we can write the completed derivative $\overline{\phi'}(x)$ as

$$\overline{\phi'}(x) = \phi'(x) + \phi(0^+)\delta(x) - \phi(L^-)\delta(x - L)\,. \tag{9b}$$

b) If $\phi: \mathbb{R} \to \mathbb{C}$ is continuous except at $x = a$ and $x = b$ where ϕ has jump discontinuities, we can calculate straightforwardly $\int_a^b \phi'(x)\, dx = \phi(b^-) - \phi(a^+)$. There is no need to introduce δ-functions since we are not integrating across the discontinuity.

c) When $\phi(x)$ is discontinuous at $x = c$, we take

$$\int_{-\infty}^{\infty} \phi(x)\delta(x - c) = \frac{1}{2}\big(\phi(c^+) + \phi(c^-)\big)$$

since $\delta(x)$ is an even function. This can lead to contradictions (Griffiths and Walborn [9]); however, identical terms of this type will appear in the expressions for both $\langle Pf|g\rangle$ and $\langle f|Pg\rangle$, so the specific definition adopted is immaterial when proving the equality.

Now we come to the main theorem of this section.

Theorem 3: P^n is Hermitian for all $n \geq 1$. We will give two proofs.

Proof #1: It suffices to prove that

$$\langle f|Pg\rangle = \langle Pf|g\rangle \tag{10}$$

when $f = P^l\varphi$ and $g = P^m\gamma$ for some ISW state vectors $|\varphi\rangle$ and $|\gamma\rangle$. Once this is established, by successive applications of (10) we obtain

$$\langle f|P^n g\rangle = \langle Pf|P^{n-1}g\rangle = \langle P^2 f|P^{n-2}g\rangle = \cdots = \langle P^{n-1}f|Pg\rangle = \langle P^n f|g\rangle\,.$$

Since $P^k = (-i\hbar D_x)^k$, both $f = P^l\varphi$ and $g = P^m\gamma$ may be discontinuous at $x = 0$ and $x = L$. By (9b):

$$\overline{g}'(x) = g'(x) + g(0^+)\delta(x) - g(L^-)\delta(x-L)\,. \tag{11}$$

We compute:

$$\begin{aligned}
-\frac{1}{i\hbar}\langle f|Pg\rangle &= \int_{-\infty}^{\infty} f(x)^*\overline{g}'(x)\,dx \\
&= \int_{-\infty}^{\infty} f(x)^*\big(g'(x) + g(0^+)\delta(x) - g(L^-)\delta(x-L)\big)\,dx \\
&= \int_0^L f^*g' + \left\{g(0^+)\int_{-\infty}^{\infty} f(x)^*\delta(x)\,dx\right\} - \left\{g(L^-)\int_{-\infty}^{\infty} f(x)^*\delta(x-L)\,dx\right\} \\
&= f^*g\Big|_{0^+}^{L^-} - \int_0^L f'^*g + \frac{1}{2}f(0^+)^*g(0^+) - \frac{1}{2}f(L^-)^*g(L^-) \\
&= -\int_0^L f'^*g + \frac{1}{2}f(L^-)^*g(L^-) - \frac{1}{2}f(0^+)^*g(0^+)
\end{aligned} \tag{12}$$

Notice that the inner product has three pieces: a boundary term, an integral over the well, and a term arising from the δ-functions that belong to $\overline{g'}$. Conversely:

$$\frac{1}{i\hbar}\langle Pf|g\rangle = \int_{-\infty}^{\infty}\big(f'(x) + f(0^+)\delta(x) - f(L^-)\delta(x-L)\big)^* g(x)\,dx$$

$$= \int_0^L f'^* g + \left\{f(0^+)^* \int_{-\infty}^{\infty}\delta(x)^* g(x)\,dx\right\} - \left\{f(L^-)^* \int_{-\infty}^{\infty}\delta(x-L)^* g(x)\,dx\right\}$$

$$= \int_0^L f'^* g + \frac{1}{2}f(0^+)^* g(0^+) - \frac{1}{2}f(L^-)^* g(L^-) \tag{13}$$

Adding (12) plus (13):

$$\langle f|Pg\rangle - \langle Pf|g\rangle = 0$$

which completes the proof.

It is instructive to show $\langle f|P^n g\rangle = \langle P^n f|g\rangle$ in a single step to illustrate the role of δ-functions in higher derivatives.

Proof #2: Suppose $|f\rangle$ and $|g\rangle$ are state vectors in the ISW. Then $f'(x)$ and $g'(x)$ are discontinuous at $x = 0$ and $x = L$. For the sake of simplicity, assume that $f'(x)$ is discontinuous only at $x = 0$. (The discontinuity at $x = L$ will be included later.) By (9a) we can write

$$\overline{f''}(x) = f''(x) + f'(0^+)\delta(x)\,.$$

Since f'' may be discontinuous at $x = 0$, we also have

$$\overline{f^{(3)}}(x) = f^{(3)}(x) + f''(0^+)\delta(x) + f'(0^+)\delta'(x)\,.$$

Continuing in this fashion:

$$\overline{f^{(n)}}(x) = f^{(n)}(x) + \sum_{k=0}^{n-1} f^{(n-k-1)}(0^+)\delta^{(k)}(x)$$

$$= f^{(n)}(x) + \sum_{k=0}^{n-1} f^{(n-k-1)}(0^+)\delta(x)(-D_x)^k \tag{14}$$

where we have used $\delta^{(k)}(x) = \delta(x)(-D_x)^k$.[k] (See [10], Example 1.29.) Therefore:

$$
\begin{aligned}
(i\hbar)^{-n}\langle P^n f|g\rangle &= \langle \overline{f^{(n)}}|g\rangle = \langle f^{(n)}|g\rangle + \sum_{k=0}^{n-1}\langle f^{(n-k-1)}(0^+)\delta(x)(-D_x)^k|g\rangle \\
&= \int_0^L f^{(n)^*} g + \sum_{k=0}^{n-1}(-1)^k f^{(n-k-1)}(0^+)^* \int_{-\infty}^{\infty}\delta(x) g^{(k)}(x)\,dx \\
&= \int_0^L f^{(n)^*} g + \frac{1}{2}\sum_{k=0}^{n-1}(-1)^k f^{(n-k-1)}(0^+)^* g^{(k)}(0^+)\,.
\end{aligned}
$$

If we include the discontinuity at $x = L$ we obtain:

$$
\begin{aligned}
(i\hbar)^{-n}\langle P^n f|g\rangle &= \int_0^L f^{(n)^*} g + \sum_{k=0}^{n-1}(-1)^k\left[f^{(n-k-1)}(0^+)^* g^{(k)}(0^+) - f^{(n-k-1)}(L^-)^* g^{(k)}(L^-)\right] \\
&= \int_0^L f^{(n)^*} g - \frac{1}{2}\sum_{k=0}^{n-1}(-1)^k f^{(n-k-1)^*} g^{(k)}\Big|_{0^+}^{L^-}\,. \qquad (15)
\end{aligned}
$$

Integration by parts can be used repeatedly on the first term:

$$
\begin{aligned}
\int_0^L f^{(n)^*} g &= f^{(n-1)^*} g\Big|_{0^+}^{L^-} - \int_0^L f^{(n-1)^*} g' \\
&= f^{(n-1)^*} g\Big|_{0^+}^{L^-} - f^{(n-2)^*} g'\Big|_{0^+}^{L^-} + \int_0^L f^{(n-2)^*} g'' \\
&= f^{(n-1)^*} g\Big|_{0^+}^{L^-} - f^{(n-2)^*} g'\Big|_{0^+}^{L^-} + f^{(n-3)^*} g''\Big|_{0^+}^{L^-} - \int_0^L f^{(n-3)^*} g^{(3)} \\
&= \cdots \\
&= \sum_{k=0}^{n-1}(-1)^k f^{(n-k-1)^*} g^{(k)}\Big|_{0^+}^{L^-} + (-1)^n \int_0^L f^* g^{(n)}\,.
\end{aligned}
$$

Therefore (15) gives:

$$
(i\hbar)^{-n}\langle P^n f|g\rangle = (-1)^n \int_0^L f^* g^{(n)} + \frac{1}{2}\sum_{k=0}^{n-1}(-1)^k f^{(n-k-1)^*} g^{(k)}\Big|_{0^+}^{L^-} \qquad (16)
$$

[k] In other words, $\int_{-\epsilon}^{\epsilon}\delta^{(k)}(x) f(x)\,dx = (-1)^k f^{(k)}(0)$.

Conversely, repeating the same steps we obtain:

$$(-i\hbar)^{-n}\langle f|P^n g\rangle = \langle f|\overline{g^{(n)}}\rangle = \int_0^L f^* g^{(n)} - \frac{1}{2}\sum_{k=0}^{n-1}(-1)^k f^{(k)^*} g^{(n-k-1)}\Big|_{0^+}^{L^-}.$$

Re-index the last sum by setting $m = n - k - 1$:

$$\begin{aligned}(-i\hbar)^{-n}\langle f|P^n g\rangle &= \int_0^L f^* g^{(n)} - \frac{1}{2}\sum_{m=0}^{n-1}(-1)^{n-m-1} f^{(n-m-1)^*} g^{(m)}\Big|_{0^+}^{L^-} \\ &= \int_0^L f^* g^{(n)} + \frac{(-1)^n}{2}\sum_{m=0}^{n-1}(-1)^m f^{(n-m-1)^*} g^{(m)}\Big|_{0^+}^{L^-}.\end{aligned}$$

Equivalently:

$$(i\hbar)^{-n}\langle f|P^n g\rangle = (-1)^n \int_0^L f^* g^{(n)} + \frac{1}{2}\sum_{m=0}^{n-1}(-1)^m f^{(n-m-1)^*} g^{(m)}\Big|_{0^+}^{L^-}. \quad (17)$$

Comparing (16) and (17) shows $\langle f|P^n g\rangle = \langle P^n f|g\rangle$.

Conclusion

The infinite square well is usually the first problem studied after the free particle. The pedagogical soundness of this presentation, however, is open to question. The artificial constraints of the ISW invoke the need for mathematical methods whose complexity is in inverse proportion to the ISW's apparent simplicity. We demonstrated the equivalence of hermicity of momentum with integration by parts, and used the ISW problem to show how boundary conditions impact the way derivatives and integrals must be treated when computing inner products of the kind $\langle f|P^n g\rangle$. More sophisticated analyses of self-adjoint operators in Hilbert spaces (Bonneau et al. [8] and Araujo et al. [11]) reveal even deeper complexities arising from the ISW's boundary conditions. The ubiquitous infinite square well is not as simple as it seems.

References

[1] D. J. Griffiths and D. F. Schroeter, "Introduction to Quantum Mechanics", 3rd edition, Cambridge University Press (2018).

[2] R. Shankar, "Principles of Quantum Mechanics", 2nd edition, Springer (1994).

[3] J. Garrison and R. Chiao, "Quantum Optics" Oxford University Press (2008).

[4] S. A. Fulling, "Comment on 'Probability Distribution of Momenta in an Infinite Square-Well Potential", American Journal of Physics, 41, 1374-1375 (1973).

[5] Y. Q. Liang, H. Zhang and Y. X. Dardenne, "Momentum Distributions for a Particle in a Box", Journal of Chemical Education, 72, 148-151 (1995).

[6] W. Rudin, "Real and Complex Analysis", 2nd edition, McGraw-Hill (1987).

[7] M. Belloni and R. W. Robinett, "The infinite well and Dirac delta function potentials as pedagogical, mathematical and physical models in quantum mechanics", Physics Reports, 540, 25-122 (2014).

[8] G. Bonneau, J. Faraut and G. Valent, "Self-adjoint extensions of operators and the teaching of quantum mechanics", American Journal of Physics, 69, 322-331 (2001).

[9] D. Griffiths and S. Walborn, "Dirac deltas and discontinuous functions", American Journal of Physics, 67, 446-447 (1999).

[10] D. H. Griffel, "Applied Functional Analysis", Ellis Horwood Ltd. (1981).

[11] V. S. Araujo, F. A. B Coutinho and J. F. Perez, "Operator domains of self-adjoint operators," American Journal of Physics, 72, 203-213 (2004).