Fonksiyonel Analiz'den Seçme Konular ve Uygulamalar

Kemal Ilgar Eroğlu (İstanbul Bilgi Üniversitesi)

Bu metin, 30 Ağustos–3 Eylül 2021 tarihlerinde Çakılarası Matematik Köyü'nde verdiğim aynı başlıklı dersin notlarından oluşmaktadır. Bu notlarda bahsedilen konular için bir kaynakça metnin sonunda verilmiştir. Bana bu olanağı veren Şahin Hoca'ya ve katılımlarıyla dersi zenginlenştiren tüm dinleyicilere teşekkürü borç bilirim.

1 Fonksiyonel Analiz'den önbilgiler

Bu ders, özellikle Kuantum Mekaniği'nin temellerini anlamak için gerekli olan matematiksel altyapıyı özetleme amacıyla hazırlanmıştır. Görüleceği üzere, bu yolda genel olarak Analiz'den, özel olarak da Fonksiyonel Analiz'den pek çok kavram karşımıza çıkacaktır.

Dersin ana konusu Hilbert uzaylarındaki dönüşümler olduğu için bu kısımda özellikle Hilbert uzaylarıyla ilgili temel tanım ve sonuçlara değineceğiz. Ancak yeri geldikçe kimi kavramların, daha genel olan norm uzayları bağlamında sunulduğu da olacaktır.

Tanım 1.1. Bir kompleks Hilbert uzayı, üzerinde ($\mathbb{C}$-değerli) bir iç çarpım tanımlı ve bu iç çarpımdan gelen norma göre (metrik anlamda) tam olan bir vektör uzayıdır. İç çarpımı $\langle \cdot, \cdot \rangle$ ile göstereceğiz ve özellikleri şunlar olacaktır: $x, y, z \in H$ ve $\alpha, \beta \in \mathbb{C}$ olmak üzere

- $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$,
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$,
- $\langle x, x \rangle \geq 0$ ve $= 0$ ancak ve ancak $x = 0$ ise.

Burada iç çarpımda gelen norm $\|x\| := \sqrt{\langle x, x \rangle}$ olarak tanımlanır. Bir Hilbert uzayında $\langle x, y \rangle = 0$ koşulunu sağlayan vektörlere birbirine dik (ortogonal) vektörler denir.

Dikkat edilirse iç çarpım ikinci bileşende konjuge-doğrusaldır:

$$\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \, \langle x, y \rangle + \overline{\beta} \, \langle x, z \rangle \, .$$

Bundan böyle aksi belirtilmedikçe tüm norm uzayları kompleks varsayılacak ve H bir kompleks Hilbert uzayını gösterecektir.

Hilbert uzayları aslında Kartezyen uzayların ($\mathbb{R}^n$ ve $\mathbb{C}^n$) sonsuz boyuta en doğal genişletmesi olarak düşünebilir. Bu uzaylardaki geometri pek çok açıdan Öklit Geometrisi'ne benzer.

Örnek 1.2. Standart iç çarpımı ile $\mathbb{C}^n$, bir kompleks Hilbert uzayıdır:

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Örnek 1.3. Önceki örneği sonsuz boyuta genişletmek istersek ℓ^2 uzayını elde ederiz. Bu uzay, $\sum_{i=1}^{\infty} |x_i|^2 < \infty$ şartını sağlayan $x = (x_i)$ kompleks sayı dizilerinden oluşur ve iç çarpım

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$$

ile verilir. Bu uzayın gerçekten bir vektör uzayı olduğu ve üstteki toplamın yakınsadığı Minkowski ve Hölder (Cauchy-Schwarz) eşitsizliklerinden görülebilir.

Örnek 1.4. $I \subseteq \mathbb{R}$ bir aralık olmak üzere bu aralıkta Lebesgue ölçülebilir olup

$$\int_I |f|^2 < \infty$$

koşulunu sağlayan $f : I \to \mathbb{C}$ fonksiyonlarını düşünelim. Bu fonksiyonların

$$f \sim g :\iff \mathcal{L}\{x : f(x) \neq g(x)\} = 0$$

yani hemen her yerde eşitlik ile tanımlanan denklik bağıntısına göre denklik sınıfları $L^2(I)$ uzayını oluşturur. Bu uzay bilindik toplama ve skaler çarpım ile bir vektör uzayıdır ve

$$\langle f, g \rangle = \int_I f \overline{g}$$

ile bir kompleks Hilbert uzayı olur.

Burada tam ifadesini vermeyeceğimiz bir teorem bize n boyutlu tüm kompleks Hilbert uzaylarının $\mathbb{C}^n$ uzayına, sonsuz boyutlu ve ayrılabilir tüm kompleks Hilbert uzaylarının da ℓ^2'ye eşyapılı olduğunu söyler. I dejenere olmayan bir aralıksa (ve Lebesgue ölçüsü kullanılıyorsa) $L^2(I)$ da sonsuz boyutlu ve ayrılabilirdir. Görüntü olarak ℓ^2'dan çok farklı dursa da aslında Hilbert uzayı olarak onunla birebir aynıdır. Nitekim, $L^2[0, 2\pi]$ uzayındaki fonksiyonların $\cos(kx)$ ve $\sin(kx)$, $k \in \mathbb{N}$ fonksiyonlarının serisi olarak açıldığı Fourier serileri bu bağlantıya iyi bir örnektir. Bu seri açılımındaki katsayılar aslında o fonksiyonu tek olarak belirleyen bir ℓ^2 dizisi verirler ve her ℓ^2 dizisi bu şekilde bir fonksiyon belirler.

Teorem 1.5 (Schwarz Eşitsizliği)**.** *Her $x, y \in H$ için*

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

1.1 Doğrusal dönüşümler

Tanım 1.6. Aynı cisim üzerinde tanımlı vektör uzayları X, Y arasında bir $T : X \to Y$ dönüşümü

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty$$

koşulunu sağlıyorsa T'ye bir doğrusal dönüşüm veya (doğrusal) operatör denir.

Eğer X ve Y norm uzayları ise ortada birer topoloji vardır ve sürekli doğrusal dönüşümlerden bahsedilebilir:

Tanım 1.7. $T : X \to Y$ iki norm uzayı arasında doğrusal dönüşüm ise T süreklidir ancak ve ancak

$$\exists C \quad \forall x \in X \quad \|Tx\| \leq C\|x\|. \tag{*}$$

Bu durumda (*) koşulunu sağlayan bir minimal $C \geq 0$ vardır ve bu değere T operatörünün normu denilir, $\|T\|$ ile gösterilir. Dolayısıyla her x için $\|Tx\| \leq \|T\|\|x\|$ sağlanır.

Aşağıdaki kısımda X ve Y norm uzayı olacaktır.

Teorem 1.8. *Norm uzaylarında bir $T : X \to Y$ doğrusal dönüşümü için şunlar eşdeğerdir:*

- *T süreklidir,*
- *T herhangi bir noktada süreklidir,*
- *T orijinde süreklidir,*
- *T düzgün süreklidir.*

Ayrıca

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|.$$

Yukarıda süreklilik $\|T\| < \infty$ olarak yorumlanacaktır.

Dikkat edilirse $\frac{\|Tx\|}{\|x\|} = \frac{\|Tx - T0\|}{\|x - 0\|}$ ifadesini T'nin grafında orijinde x yönündeki eğim olarak yorumlayabiliriz. Sonlu boyutta yönler kümesi kompakt olduğu için eğimlerin sınırlı olduğu sezilebilir, nitekim gerçekten de $\dim X < \infty$ ise T sürekli olmak zorundadır. Ancak $\dim X = \infty$ ise eğimler üstten sınırlı olmayabilir, bu da süreksizlik demektir.

Operatör normunun özellikleri:

- $S, T : X \to Y$ sürekli ise $\|S + T\| \leq \|S\| + \|T\|$,
- $\|\alpha T\| = |\alpha|\|T\|$,
- $\|T\| \geq 0$ ve $= 0$ ancak ve ancak $T = 0$ ise,
- $ST := S \circ T$ olmak üzere $\|ST\| \leq \|S\|\|T\|$,
- $T : X \to X$ ise bir üstteki özellikten her $n \in \mathbb{N}$ için $\|T^n\| \leq \|T\|^n$ gelir.

Bunlardan sonra şu sonuç şaşırtıcı olmaz:

Teorem 1.9. *$B(X, Y)$ ile sürekli $T : X \to Y$ doğrusal dönüşümlerinin kümesini gösterelim. Bilindik işlemlerle bu uzay bir vektör uzayıdır. Üstelik, operatör normu bu uzayda bir normdur ve bu norm altında $B(X, Y)$ tamdır (Banach uzayıdır) ancak ve ancak Y tam (Banach) ise.*

1.2 Sürekli doğrusal dönüşümlerin fonksiyonları

Şimdi X bir tam kompleks norm uzayı (yani kompleks Banach uzayı) ve $T : X \to X$ bir sürekli doğrusal dönüşüm olsun. Elbette kompleks katsayılı her p polinomu için $p(T)$ doğrusal dönüşümünü tanımlayabiliriz. Peki T'nin bir kuvvet serisinden bahsedilebilir mi?

En basit örnek olarak $\|T\| < 1$ durumunu ele alalım. Kompleks sayılarda $|z| < 1$ için

$$\frac{1}{1-z} = (1-z)^{-1} = 1 + z + z^2 + z^3 + \cdots = \sum_{n=0}^{\infty} z^n$$

özdeşliğinin nasıl kanıtlandığını hatırlayınız. Şimdi $B(X) := B(X, X)$ uzayının operatör normu altında tam olmasını kullanarak (Teorem 1.9), aynı kanıtın z yerine T alınarak birebir şekilde operatörlere uyarlanabileceğini gözleyiniz. Yani $\|T\| < 1$ ise $1 - T := I - T$ operatörünün X'ten X'e sürekli bir ters dönüşümü vardır ve bu dönüşüm

$$\frac{1}{1-T} := (1-T)^{-1} = \sum_{n=0}^{\infty} T^n \tag{1}$$

ile verilir. Sağdaki seri kısmi toplamları $B(X)$'te (mutlak) yakınsak bir seri olacaktır. Kanıt uyarlanırken görülecek bir ufak fark, sayılar için $|z^n| = |z|^n$ iken operatörler için $\|T^n\| \leq \|T\|^n$ olmasıdır; ama bu değişiklik kanıtın beklendiği gibi yürümesini engellemez. Genel olarak, $\lambda \in \mathbb{C}$ ve $\|T\| < \lambda$ ise, $\|T/\lambda\| < 1$ olmasını kullanarak

$$\frac{1}{\lambda - T} := (\lambda I - T)^{-1} = \frac{1}{\lambda} \cdot \frac{1}{1 - (T/\lambda)} = \sum_{n-0}^{\infty} \lambda^{-n-1} T^n \tag{2}$$

elde edilir.

Benzer şekilde $z_0 \in \mathbb{C}$ merkezli ve $R > 0$ yakınsaklık yarıçaplı

$$f(z) := \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

kuvvet serisini düşünelim. Bu durumda herhangi bir $T_0 \in B(X)$ verildiğinde, $\|T - T_0\| < R$ koşulunu sağlayan her $T \in B(X)$ için

$$f(T) = \sum_{n=0}^{\infty} a_n (T - T_0)^n$$

toplamının anlamlı olduğu, bu yolla böyle T'ler için $f(T)$ fonksiyonunun tanımlanabileceği görülebilir. İleride bu konuya tekrar döneceğiz.

1.3 İzdüşümler ve fonksiyoneller

Teorem 1.10. *H Hilbert ve $\emptyset \neq M \subseteq H$ kapalı ve konveks bir altküme olsun. Her $x \in H$ için M içinde x'e en yakın biricik bir y_0 noktası vardır; yani*

$$\forall x \in H \quad \exists! y_0 \in M \quad \|x - y_0\| = dist(x, M) := \inf_{y \in M} \|x - y\|.$$

Bu en yakın noktaya x'in M'ye izdüşümü denir. Eğer özel olarak M kapalı bir altuzaysa, buna dik (ortogonal) izdüşüm denir ve y_0 noktasını karakterize eden özellik $x - y_0 \perp M$ olmasıdır, yani her $y \in M$ için $\langle x - y_0, y\rangle = 0$.

Bir vektörün bir başka vektör yönündeki izdüşümünden kasıt, ikinci vektörün gerdiği bir boyutlu altuzaya izdüşümüdür. Şu bilgi çok kullanışlıdır:

Teorem 1.11. *$x \in H$ herhangi bir vektör ve $e \in H$ bir birim vektör olsun. Bu durumda x'in e üzerine izdüşümü $\langle x, e\rangle e$ vektörüdür.*

Bir X vektör uzayından skaler cisme giden dönüşümlere *fonksiyonel* denir. Eğer $f : X \to \mathbb{C}$ bir doğrusal fonksiyonel ise elbette

$$Ker(f) := \{x : f(x) = 0\}$$

çekirdek uzayı X'te bir altuzaydır ve $f \neq 0$ ise tümleyen boyutu (codimension) 1 olan bir özaltuzaydır. Dolayısıyla $f = const$ kümeleri aslında $Ker(f)$ altuzayının ötelemelerinden oluşan ve birbirine paralel afin hiperdüzelemlerdir. Geometrik olarak bir x noktası için $f(x)$ değerini belirleyen şey, x'i içeren hiperdüzlemin $Ker(f)$'in ne kadar ötelenmesiyle elde edildiğini belirten parametredir. Tanım uzayı bir Hilbert uzayı ise bu ötelemenin miktarı, $Ker(f)$ altuzayına dik bir vektöre alınacak izdüşümle hesaplanabilir. Bu gözlem, aşağıdaki teoremin kanıtının ana fikridir:

Teorem 1.12 (Riesz Temsil Teoremi)**.** *Her sürekli $f : H \to \mathbb{C}$ doğrusal fonksiyoneli için*

$$\forall x \in H \qquad f(x) = \langle x, z\rangle$$

koşulunu sağlayan biricik bir $z = z(f) =: z_f \in H$ elemanı vardır. Üstelik bu durumda $\|z_f\| = \|f\|$ olur ve $f \to z_f$ eşlemesi, sürekli doğrusal f'ler uzayından ("H'nin dual uzayı") H'ye konjuge-doğrusal bir izomorfizmadır.

1.4 Hilbert eşlenik operatörü

Şimdi $T : H \to H$ bir sürekli doğrusal dönüşüm olsun. Bir $y \in H$ alalım ve $f_y : H \to \mathbb{C}$ dönüşümünü

$$f_y(x) = \langle Tx, y\rangle$$

ile tanımlayalım. Schwarz eşitsizliğine göre f_y süreklidir:

$$|f_y(x)| = |\langle Tx, y\rangle| \leq \|Tx\|\|y\| \leq \|T\|\|x\|\|y\|$$

(yani $C = \|T\|\|y\|$ alınabilir). Bu durumda Riesz Temsil Teoremi bize her $x \in H$ için

$$f_y(x) = \langle Tx, y\rangle = \langle x, y^*\rangle$$

şartını sağlayan biricik bir $y^* \in H$ elemanının varlığını söyler. $T^* : H \to H$ dönüşümünü

$$T^* y := y^*$$

ile tanımlayalım, yani her x, y için

$$\langle Tx, y\rangle = \langle x, T^* y\rangle .$$

Elbette eşlenik alarak $\langle x, Ty\rangle = \langle T^* x, y\rangle$ olduğu da görülür. Yani kuralın işleyişini görsel olarak "T karşıya bir yıldız alarak geçer" şeklinde özetleyebiliriz.

Tanım 1.13. Bu T^* dönüşümüne T'nin Hilbert eşleniği (Hilbert adjoint) denir.

Teorem 1.14. *T^* doğrusaldır ve*

- $\|T^*\| = \|T\|$,
- $(\alpha S + \beta T)^* = \overline{\alpha} S^* + \overline{\beta} T^*$,
- $(T^*)^* = T$,
- $(ST)^* = T^* S^*$,
- *Eğer $T^{-1} : H \to H$ varsa süreklidir ve*

$$(T^{-1})^* = (T^*)^{-1}.$$

Fonksiyonel analizdeki pek çok önemli operatör sınıfı Hilbert eşleniği üzerinden tanımlanır. Bu dersin ana konusu olan özeşlenik dönüşümler de bunlardan biridir:

Tanım 1.15. $T : H \to H$ sürekli olsun. Eğer
(i) $TT^* = T^*T$ ise T'ye *normal*,
(ii) $T^{-1} = T^*$ ise T'ye *üniter* (unitary),
(iii) $T = T^*$ ise T'ye *özeşlenik* (Hermityen, self-adjoint) dönüşümdür denir.

Eğer $T : \mathbb{C}^n \to \mathbb{C}^n$ ve T'yi standart tabanda temsil eden $n \times n$ matris A ise, T^*'ı temsil eden matris $A^* := \overline{A}^T$ matrisidir. Yukarıdaki tanımlar T yerine A yazılarak matrislere de uygulanır.

Teorem 1.16. *$T : H \to H$ bir doğrusal dönüşüm ise, T sürekli ve özeşleniktir ancak ve ancak T simetrik ise, yani her $x, y \in H$ için*

$$\langle Tx, y\rangle = \langle x, Ty\rangle$$

ise.

Özeşlenikliğin simetriyi getirmesi apaçıktır. Tersi yönde ise T'nin sürekliliği elde edildikten sonra özeşleniklik hemen çıkar. Süreklilik ise $\mathcal{F} = \{f_y(\cdot) = \langle \cdot, Ty\rangle : \|y\| = 1\}$ sürekli fonksiyonel kümesine aşağıda verilen Düzgün Sınırlılık İlkesi uygulanarak ve $\|f_y\| = \|Ty\|$ olduğu hatırlanarak elde edilebilir:

Teorem 1.17 (Düzgün Sınırlılık İlkesi)**.** *X ve Y norm uzayları, X tam (Banach) olsun. X'ten Y'ye giden sürekli doğrusal dönüşümlerden oluşan bir $\mathcal{F} = \{T_\alpha : \alpha \in I\}$ kümesi alalım. Eğer her $x \in X$ için*

$$S_x := \{T_\alpha x : \alpha \in I\} \subseteq Y$$

kümesi Y'de sınırlı ise, $\mathcal{F}$ kümesi $B(X, Y)$'de sınırlıdır, yani $\sup_{\alpha \in I} \|T_\alpha\| < \infty$.

Örnek 1.18. Bir kompakt $[a,b]$ aralığı için $H = L^2[a,b]$ olsun ve $T : H \to H$

$$(Tx)(q) := qx(q)$$

ile tanımlansın. Bu durumda T açıkça doğrusaldır. Üstelik simetriden ötürü özeşleniktir:

$$\langle Tx, y \rangle = \int_a^b qx(q)\overline{y(q)}dq = \int_a^b x(q)\overline{qy(q)}dq = \langle x, Ty \rangle .$$

Burada T'nin sürekliliğini doğrudan görmek de kolaydır: $C = \max\{|a|,|b|\}$olmak üzere

$$\begin{aligned}\|Tx\|^2 = \langle Tx, Tx \rangle &= \int_a^b q^2|x(q)|^2 dq \\ &\leq C^2 \int_a^b |x(q)|^2 dq = C^2\|x\|^2.\end{aligned}$$

1.5 Özdeğerler, özvektörler

X bir kompleks vektör uzayı ve $T : X \to X$ bir doğrusal dönüşüm olsun.

Tanım 1.19. Eğer bir $\lambda \in \mathbb{C}$ için $Tx = \lambda x$ şartını sağlayan bir $x \neq 0$ varsa, λ sayısı T'nin bir *özdeğeridir* (eigenvalue) ve x de T'nin bir λ-*özvektörüdür* (eigenvector) denir. Bu durumda λ-özvektörlere 0'ı ekleyerek elde edilen $E_\lambda := Ker(T-\lambda) := Ker(T-\lambda I)$ altuzayına da λ-özvektör altuzayı denir (eigenspace).

Örnek 1.20. $\dim X = n < \infty$ ve T belli bir tabanda A matrisi ile temsil ediliyorsa, λ bir özdeğerdir ancak ve ancak $Ker(T-\lambda) \neq \{0\}$ ise, ki bu da $\lambda - A$ matrisinin tekil (singüler) olması yani

$$p(\lambda) := \det(\lambda - A) = 0$$

olmasıyla denktir. Üstteki $p(\lambda)$'ya A'nin (veya T'nin) karakteristik polinomu denir ve bu, derecesi n olan monik bir polinomdur. Demek ki özdeğerler tamı tamına karakteristik polinomun kökleridir. Özvektörler matris temsili üzerinden kolayca hesaplanabilir.

Bu noktada kompleks vektör uzaylarında çalışmamızın en önemli gerekçelerinden birinin kompleks cismin tamlığı sayesinde üstteki senaryoda daima (katlılıklarıyla birlikte) n tane özdeğer bulunabilmesi olduğunu belirtmek isteriz. Özellikle T'nin özeşlenik olduğu durumda, özvektörlerden bir taban elde etmek mümkün olmaktadır ve bu çok ideal bir durumdur. Sonsuz boyuttaki dönüşümlerde de geçerli olan başka kolaylıklar da vardır:

Teorem 1.21. *$T : H \to H$ özeşlenik ise*

(a) T'nin özdeğerleri gerçeldir.

(b) Farklı özdeğerlerin özvektörleri birbirine diktir.

(c) $\dim H < \infty$ ise H'nin T-özvektörlerinden oluşan bir ortonormal tabanı vardır (ortonormal, "birbirine dik birim vektörlerden oluşan" demektir).

İlk iki şıkkı kanıtlayalım: $0 \neq x$ bir λ-özvektör ise

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Tx, x \rangle = \langle x, Tx \rangle = \langle x, \lambda x \rangle = \overline{\lambda} \langle x, x \rangle$$

ve buradan $\lambda = \overline{\lambda}$ gelir, yani $\lambda \in \mathbb{R}$. Şimdi $\lambda_1 \neq \lambda_2$ birer özdeğer ve x_1, x_2 de sırayla bunlara ait özvektör ise

$$\begin{aligned}\lambda_1 \langle x_1, x_2 \rangle &= \langle Tx_1, x_2 \rangle = \langle x_1, Tx_2 \rangle \\ &= \overline{\lambda_2} \langle x_1, x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle\end{aligned}$$

olmasından $\langle x_1, x_2 \rangle = 0$ gelir; burada $\lambda_2 \in \mathbb{R}$ olmasını kullandık.

Bu son teoremden yola çıkarak $\dim H = n$ olmak üzere elimizde bir $T : H \to H$ özeşlenik dönüşümü ve bu dönüşümün özvektörlerinden oluşan bir $E = (e_1, \ldots, e_n)$ ortonormal tabanı olsun ve $Te_i = \lambda_i e_i$ yazalım. Bu durumda her x vektörü

$$x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i$$

şeklinde ortogonal bileşenlerine ayrılabilir. Burada E tabanı Kartezyen uzaydaki standart tabana benzer bir işlev görür ve $\langle x, e_i \rangle$ katsayıları da Kartezyen koordinatların yerini alır.

Bu tür bir temsil sonsuz boyutta da geçerlidir. Örneğin elimizde T-özvektörlerden oluşan bir $(e_i)_{i=1}^{\infty}$ ortonormal dizisi olsa ve bu dizi tam (complete, total) bir küme oluştursa, yani gerdiği altuzay H'de yoğun olsa; eşdeğer söylemle, tüm e_i'lere dik tek vektör 0 vektörü olsa (bu da e_i'lerin H'deki tüm "yönleri" üretecek kadar zengin bir küme olduğu anlamına gelir), bu durumda her $x \in H$ için biricik bir

$$x = \sum_{i=1}^{\infty} \alpha_i e_i$$

temsili vardır ve bu da elbette

$$x = \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \tag{3}$$

ile verilendir. Buna x'in bir Fourier serisi denir. Buradaki e_i'ler klasik anlamda H'nin bir (Hamel) tabanını oluşturmaz ama "sonsuz lineer kombinasyon"larla tüm vektörleri üretirler. Burada x'i sonsuz sayıdaki dik kenarına parçalanmış bir hipotenüs olarak düşünebiliriz ve Parseval Özdeşliği denilen

$$\|x\|^2 = \sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \tag{4}$$

eşitliği geçerlidir. Bu aslında Pisagor Kanunu'nun sonsuz boyutlu halidir. Üstte olduğu gibi $Te_i = \lambda_i e_i$ ise bu durumda

$$Tx = \sum_{i=1}^{\infty} \langle x, e_i \rangle Te_i = \sum_{i=1}^{\infty} \lambda_i \langle x, e_i \rangle e_i$$

şeklinde açılabilecektir.

Peki acaba T'nin özdeğerleri bir $I \subseteq \mathbb{R}$ aralığından gelebilir mi? Bu durumda ortonormal tabanımız $(e_\lambda)_{\lambda\in\mathbb{R}}$ olursa artık toplamların integrale dönüştüğü

$$x = \int_I \langle x, e_\lambda \rangle \, e_\lambda d\lambda \quad \text{ve} \quad Tx = \int_I \lambda \, \langle x, e_\lambda \rangle \, e_\lambda d\lambda \qquad (5)$$

benzeri temsiller oluşturabilir miyiz? Bu, anlamlı olur mu?

Aslında tüm ayrılabilir Hilbert uzaylarında her ortonormal küme sayılabilir çoklukta olmak zorundadır ve sonuç olarak en fazla sayılabilir tane özdeğer olabilir. Ama üstteki sezgi de tamamen boş bir yanılsamadan ibaret değildir ve bizi fizikçilerin de kullandığı Fourier integrallerine ve Fourier dönüşümüne götürür. Bir sonraki kısımda fizikçilerin Kuantum Mekaniği hesaplarında özeşlenik dönüşümlere ait nasıl sayılamaz çoklukta "özdeğer" bulduğunu ve bunları nasıl kullandıklarını örnekleyeceğiz. Bu dersin temel amaçlarından biri bu sezgisel yaklaşımın nasıl matematiksel olarak sağlam zemine oturtulduğunun açıklanmasıdır. Bunu ileriki bölümlerdeki spektrum kavramı ve von Neumann'ın temsil teoremi yardımıyla yapacağız.

2 Kuantum Mekaniği'ndeki tablo

Kuantum Mekaniği'nin temelleri 1900'lerin ilk çeyreğinde atılmıştır. Ondokuzuncu yüzyılın sonunda zaten Klasik Mekanik'in eksikliği fark edilmiş ve (yaygın kabul görmesi uzun sürse de) Einstein'ın Genel Görelilik Kuramı makro ölçekteki sorulara bir yanıt getirmeyi başarmıştı. Ancak buna koşut olarak atom ölçeğinde de bazı "tuhaf" davranışlar gözlenmekteydi ve bu olgular hâlâ açıklanmaya muhtaçtı.

Örneğin kara cisim ışımasındaki enerji problemini Planck ancak enerjinin sürekli bir yapıda değil de paketler halinde yayıldığını varsayarak çözebilmişti. Çift yarık deneyinde madde parçacıkları tanecikten çok dalga gibi davranıyordu. Atomun pozitif yüklü çekirdek etrafında dönen negatif yüklü elektronlardan oluştuğu fikri öne çıkmıştı ama görünüşe göre bu elektronlar Güneş etrafındaki gezegenler gibi herhangi bir uzaklıkta (ya da enerji düzeyinde) değil ancak kesikli değerler alan belli düzeylerde bulunabiliyordu. Bu, klasik elektromanyetik kuramla açıklanabilir bir durum değildi.

Bu kesikli yapıyı açıklama yönündeki ilk girişimlerden biri 1910'larda ortaya çıkan Bohr-Wilson-Sommerfeld Kuantizasyonu idi ancak bu da istenen ölçüde yeterli olmadı. 1920'lerde ise de Broglie'nin madde parçacıklarının da bir dalga olarak ele alınması önerisinin hemen ardından 1925'te Heisenberg'in Bohr atom modelini açıklamak için öne sürdüğü bir "matris formülasyonu" ortaya çıktı. Nihayet 1926'da Schrödinger bir dizi makale ile Heisenberg'in modelini (sonsuz boyutlu) Hilbert uzayları üzerinde bir özeşlenik operatörler kuramına genişletti. Kuantum Mekaniği günümüzde yerini Kuantum Alan Kuramı'na bırakmış olsa da, Klasik Mekanik'in yeterli olduğu problemlerde hâlâ kullanılıyor olması gibi, Kuantum Mekaniği de belli olguları açıklamak için kullanılmaktadır ve bugünkü kuram Schrödinger'in öne sürdüğü temel üzerine kurulmuştur.

Elbette Schrödinger'in modeli deneysel açıdan çok tatmin ediciyse de, matematiksel temeli açısından boşluklarla doluydu. Bu boşlukların kapatılması işi 1920'lerden 40'lara kadar uzanan bir süreçte Fonksiyonel Analiz'in gelişmesindeki en büyük itici güçlerden biri oldu. Von Neumann, Stone, Weyl gibi büyük isimler bu yoldaki başrol oyuncular oldular.

Bugün matematiksel temel açısından bir eksiklik kalmamış olsa da Kuantum Mekaniği kitaplarının çoğunda anlatımın eski yöntemle yapılmaya devam ettiğini görmekteyiz. Bu nedenle bu metinler bir matematikçi gözünde anlaşılmaz ve hatta düpedüz yanlış gözükebilmektedir. İşte Çakılarası'nda verilen bu ders dizisinin bir amacı da, fizik kitaplarının anlatmadığı matematiksel arkaplan hakkında temel bilgileri vermektir.

Altta verilecek "fizikçinin bakışı"na örnekler için Vladimir Fok'un kitabındaki [5] anlatım temel alınmıştır. Daha matematiksel bir bakış için Faddeev-Yakubovskiĭ [4], ya da en ayrıntılı anlatım için von Neumann [6] veya Takhtajan'a [7] bakılabilir.

Şimdi Kuantum Mekaniği'nin matematiksel araçlarını oluştururken izleyeceğimiz temel ilkeleri özetleyelim. Modelimiz, bir "kuantum sistemi"nin ve bu sistem üzerinde yapılacak ölçümlerin modellenmesiyle ilgilidir. Bu bağlamda kuantum sisteminden anlayacağımız tek bir elektron, ya da bir bütün olarak tanecik gözüyle bakılan bir atom gibi, küçük ölçekli çok basit bir sistem olacaktır. Birden çok tanecikli sistemlerin modellenmesi hakkında ileride birkaç yorumda bulunacağız (Kısım 7.3).

Temel ilkelerimiz şunlardır:

(a) Bir fiziksel sistemin durumu (hal, state) bir Hilbert uzayına ait bir vektör ile temsil edilecektir.

(b) Bazı fiziksel büyüklükler aynı anda ölçülemeyebilir.

(c) Ölçülen büyüklüklerin olası değerleri tüm gerçel sayılara dağılmışsa (konum, momentum vb.) ilk maddedeki vektör, aynı anda ölçülebilen fiziksel büyüklüklerin bir fonksiyonu olarak verilen bir ψ "dalga fonksiyonu" olarak seçilebilir. Bu durumda Hilbert uzayımız $L^2(\mathbb{R}^n)$ olacaktır. Burada n, fonksiyonun parametre sayısına göre değişebilir; bu sayı da sistemin serbestlik derecesiyle ilgilidir. Dolayısıyla üç boyutlu problemlerde $n = 3$ alınması yeterlidir.

(d) Bir fiziksel büyüklüğün ölçümü sistemin durumunu değiştirir. Bu değişim, ψ üzerine bir özeşlenik doğrusal dönüşüm şeklinde etkir. Dolayısıyla ölçülebilen her fiziksel büyüklük için o ölçümü temsil eden bir özeşlenik dönüşüm (operatör) vardır.

(e) Ölçüm sonucu bu özeşlenik dönüşümün bir özdeğeri[1] çıkmak zorundadır (ölçüm sonucunun gerçel sayı çıkması gerekir, bu da doğal olarak bizi gerçel özdeğerlere sahip dönüşümleri kullanmaya iter. Bu ve ileride göreceğimiz başka gerekçelerle en uygun aday özeşlenik dönüşümlerdir).

(f) Bir a fiziksel büyüklüğünün ölçümü A operatörü ile temsil ediliyorsa ve sistemin durumunu temsil eden

[1] Aslında burada kastedilen özdeğer değil spektral değerdir. Bunu ileride göreceğiz.

dalga fonksiyonu a'nın bir fonksiyonu olarak $\psi(a)$ ile veriliyorsa, A'nin ψ üzerindeki etkisi a ile çarpımdır; yani A dönüşümü $\psi(a)$ fonksiyonunu $a\psi(a)$ fonksiyonuna taşır. Daha genel olarak, sürekli bir $f : \mathbb{R} \to \mathbb{R}$ için $f(a)$ büyüklüğünü ölçen operatör $\psi(a)$'yı $f(a)\psi(a)$'ya taşır.

Bu son maddede öne çıkan bir durumu vurgulayalım: Sistemin durumunu temsil eden dalga fonksiyonu a büyüklüğünü parametre alan bir $\psi(a)$ fonksiyonu olabileceği gibi bir başka b büyüklüğünü parametre alan $\phi(b)$ fonksiyonu da olabilir. Bu durumda A'nın ψ üzerine etkisi çarpma iken $\phi(b)$ temsili üzerindeki etkisi bambaşka olabilir. Bunun örneğini birazdan konum ve momentum operatörleri üzerinde göreceğiz.

Bir de genel olarak bize yol gösterecek olacak bir kuralımız olacak:

Karşılıklılık İlkesi: Kuantum Mekaniği'nde belli bir fiziksel büyüklüğü ölçen dönüşümler, bunların Klasik Mekanik'teki karşılıklarıyla eşleşmelidir. Bundan ne kastedildiğini ilerideki örneklerde daha iyi anlayacağız.

Birinci ve üçüncü maddede bahsi geçen Hilbert uzayının seçimi aslında sistemin serbestlik derecesi, ölçümün alabileceği değerler ve o ölçümü veren dönüşümlerin sağlaması beklenen cebirsel ilişkilere göre yapılır. Konum ve momentum için sonsuz boyutlu $L(\mathbb{R}^n)$'e ihtiyaç duyulurken belli bir yöndeki polarizasyon gibi 2 değerli bir ölçüm $\mathbb{C}^2$'deki temsillerle modellenebilir.

Örneklere geçmeden önce bir uyarı: Yukarıdaki anlatım yine Fok'un kitabındakine benzeyen ve çoğu modern Kuantum Mekaniği kitabında görülebilecek türden bir sunuşun biraz basitleştirilmiş bir özetidir. Modern matematiksel yaklaşımda ise sadece ölçümler değil, sistemin durumları da birer özeşlenik dönüşümle verilir. Sistemin halini veren özeşlenik dönüşümler iz sınıfından (trace class) olup izi (trace) 1 olan özeşlenik dönüşümlerdir. Bunlar bir konveks küme oluşturur. Yukarıdaki anlatımda ψ ile temsil edilen haller bu modelde ψ'nin gerdiği bir boyutlu altuzaya dik izdüşümü veren P_ψ dönüşümleridir, bunlara saf hal (pure state) denir. Bu izdüşümler, tamı tamına tüm haller kümesinin (konvekslik kuramındaki anlamıyla) uç noktalarıdır.

Bir de şunu unutmamakta yarar var: Schrödinger'i böyle bir modeli kullanmaya iten gerekçeleri uzun uzadıya tartışmış değiliz. Bu konuda daha ayrıntılı bir açıklama için Fok'un kitabına bakılabilir. Ama günün sonunda olan şey şuydu: Deneylerde görülen olguları açıklayacak bir matematiksel kuram gerekiyordu. Deneylerin verdiği ipuçlarını izleyen bir grup fizik dehası, sonunda özeşlenik dönüşümler üzerine kurulu bir model tahmininde bulundular ve bu tahmin sonraki deneylerin sınamasından başarıyla geçti. Üstelik onlar bu modeli kurarken Klasik Mekanik'i tümden silmek yerine, tam tersine ona olabildiğince bağlı kalan bir yol izlemeye çalıştılar. İlginçtir ki, kanıtı bu dersin dışında kalıyor olsa da, kuantum sistemlerinin "kesiklilik" birimi olarak kabul edebileceğimiz Planck sabiti $\hbar$ sıfıra giderken Kuantum Mekaniği'nin limit halinin Klasik Mekanik olduğu gösterilebilir! Yani gerçekten de $\hbar$'nin "çok küçük" kaldığı ve fiziğin kesiksiz, sürekli göründüğü ölçeklerin mekanik kuramı Klasik Mekanik'tir! Bu gözlem, tıpkı Görelilik Kuramı'nın düşük hızlarda Klasik Mekanik'e indirgenmesi gibi, Kuantum Mekaniği'nin yaklaşımının doğru yönde bir genelleme olduğuna olan güveni pekiştirir.

Şimdi örneklerimize geçelim:

Örnek 2.1. Bir boyutlu bir sistemde q konumu gösteriyorsa ve sistemin dalga fonksiyonu (yani durumu, hali) $\psi(q)$ ile veriliyorsa, konum operatörü Q

$$(Q\psi)(q) = q\psi(q)$$

ile verilir, biz bunu (doğru anlaşılmasını umarak) $Q : \psi(q) \mapsto q\psi(q)$ ile göstereceğiz. Daha yüksek boyutta ise, yani $q = (q_1, \ldots, q_n)$ ise, konumun i. bileşenini ölçen Q_i operatörü

$$\psi(q) = \psi(q_1, \ldots, q_n) \mapsto q_i\psi(q_1, \ldots, q_n)$$

ile verilir.

Örnek 2.2. Yine bir boyutta q konum ise, momentum ölçümünü veren P operatörü

$$P : \psi(q) \mapsto -i\hbar\psi'(q)$$

ile verilir. Benzer şekilde, $n \geq 2$ boyutta ise momentumun i. bileşeni

$$P_i : \psi(q_1, \ldots, q_n) \mapsto -i\hbar\frac{\partial\psi}{\partial q_i}(q_1, \ldots, q_n)$$

ile ölçülür; burada $\hbar > 0$ Planck sabitini göstermektedir (bu sabitin değeri deneylerden elde edilmektedir).

Özetle, dalga fonksiyonu konum cinsinden ise, konum operatörü çarpım, momentum operatörü ise türev formundadır. Fizik kitaplarında bu operatörler yukarıdaki gibi tanıtılıp özeşlenik oldukları belirtilir ve yola devam edilir.

Ama matematikçi gözüyle bakınca önümüzde büyük sorunlar var!

Biz sorunları Q üzerinden örnekleyeceğiz; ama P için de durum çok benzerdir.

(i) Bir kere Q, tüm $L^2(\mathbb{R})$'de tanımlı değildir! $\psi(q) \in L^2$ olup $q\psi(q) \notin L^2$ olan ψ'ler vardır, örnek olarak C^∞ sınıfından, analitik açıdan pek güzel olan

$$\psi(q) = \frac{q}{1+q^2}$$

fonksiyonu alınabilir. Öte yandan, Q'yu

$$D(Q) := \{\psi(q) \in L^2(\mathbb{R}) : q\psi(q) \in L^2(\mathbb{R})\}$$

altuzayında (bunun bir altuzay olduğunu gösteriniz!) tanımlayabiliriz. Üstelik bu L^2 içinde yoğun bir altuzaydır; bir başka yoğun altuzay olan $C_0^\infty(\mathbb{R})$'yi (kompakt destekli C^∞ fonksiyonlar) içerir.

(ii) Ne var ki $Q : D(Q) \subseteq L^2(\mathbb{R}) \to L^2(\mathbb{R})$ şeklinde iyi tanımlı olsa da, sürekli değildir! Örneğin bir boyutta $\psi_m = \chi_{[m,m+1]}$, $m \in \mathbb{N}$ fonksiyonlarına bakabiliriz.

Burada "S kümesinin karakteristik fonksiyonu" olarak isimlendirilen ve

$$\chi_S(x) = \begin{cases} 1, & x \in S \\ 0, & x \notin S \end{cases}$$

olarak tanımlanan fonksiyonu kullanmış olduk. Buna göre

$$\|\psi_m\|^2 = \int_{\mathbb{R}} |\chi_{[m,m+1]}(q)|^2 dq = \int_m^{m+1} dq = 1$$

olur ama

$$\|Q\psi_m\|^2 = \int_{\mathbb{R}} |q\chi_{[m,m+1]}(q)|^2 = \int_m^{m+1} q^2 dq \geq m^2$$

sağlanır ve $\|Q\psi_m\|/\|\psi_m\| \overset{m\to\infty}{\longrightarrow} \infty$ olmasından Q'nun süreksizliği ortaya çıkar. Hem tanım kümesinin küçüklüğü hem de süreksizlik, ilk bölümdeki Riesz Temsil Teoremi'nin hipotezlerini bozar ve Q^* eşlenik dönüşümünün varlığı tehlikeye girer. Peki bu durumda özeşleniklikten nasıl bahsederiz?

(iii) Tüm bunlar yetmezmiş gibi, Q'nun özdeğeri yoktur! Gerçekten de $Q\psi = \lambda\psi$ yani her q için

$$q\psi(q) = \lambda\psi(q)$$

şartını sağlayan tek fonksiyon ancak $\psi = 0$ olabilir. Halbuki konum ölçümünün sonucu bir özdeğer çıkmalıydı!

Kuantum Mekaniği deneylerde iyi çalıştığına göre, her şeye rağmen fizikçilerinin sezgilerinin boş olmaması gerekir. Bütün bunlar bizi özeşleniklik, özvektör vb. kavramların daha dikkatli şekilde tanımlandığı ya da genellendiği bir arayışa iter.

2.1 Fizikçiler bu durumda ne diyor?

Matematiksel temeli sağlamlaştırmadan önce fizikçilerin bu "sahte" özeşleniklik dönüşümleri ve onların "sahte" özvektörlerini nasıl kullandığını örnekleyelim. Bu kısımda yine Fok'un anlatımını izleyeceğiz ve örneklerimiz P momentum operatöründen gelecek. Yalınlık açısından bir boyutta kalacağız, ama daha yüksek boyutlardaki irdeleme benzerdir.

Elde bir özeşlenik operatör olunca fizikçinin amacı bunun özvektörlerinden oluşan (3) veya (5) anlamında bir "ortonormal taban" elde etmektir. Her bir özdeğerin özvektörleri bulunup bunlar "normalize" edilerek yani ψ yerine $\psi/\|\psi\|$ vektörüne geçilerek bir ortonormal taban elde edilmeye çalışılır. Dikkat edilirse özeşleniklikten ötürü farklı özdeğerlerin özvektörleri birbirlerine zaten dik olacaktır.

Konuma bağlı $\psi(q)$ dalga fonksiyonu üzerinden P'nin özdeğerlerini ve özvektörlerini bulmaya çalışalım. Denklemimiz $P\psi = \lambda\psi$, yani

$$\begin{aligned} -i\hbar\psi'(q) &= \lambda\psi(q) \\ \frac{d\psi}{dq} &= \frac{i\lambda}{\hbar}\psi(q) \end{aligned}$$

şeklindedir. Belli bir $\lambda \in \mathbb{R}$ için bunun çözümü elbette(!) $c(\lambda) \in \mathbb{C}$ bir sabit olmak üzere

$$\psi(\lambda, q) = \psi_\lambda(q) = c(\lambda)e^{i\lambda q/\hbar}$$

fonksiyonudur, yani her $\lambda \in \mathbb{R}$ bir özdeğer olmuş oldu ve her özdeğer için bir boyutlu (tek özvektör ile gerilen) bir özaltuzay olduğunu bulduk. Tabii ki aslında $|\psi_\lambda(q)| \equiv 1$ olduğundan $\psi_\lambda \notin L^2$, yani ortada gerçek anlamda özdeğer veya özvektör yoktur!

Fizikçi bu durumu "bu durumda λ bir has (proper) özdeğer değildir ama sürekli spektruma aittir deriz" sözleriyle geçiştirir ama ψ_λ fonksiyonlarını kullanmaktan geri kalmaz! Gerçek anlamda özdeğer olan λ değerleri için normalizasyon işlemini beklendiği gibi $\psi_\lambda \mapsto \psi_\lambda/\|\psi_\lambda\|$ geçişiyle yapar. Ama "sürekli spektrum"dan gelen özdeğerler(!) için şöyle bir özel bir normalizasyon prosedürü uygular:

Amaç, $c(\lambda)$ için uygun bir değer, tercihen bir pozitif gerçel sayı seçerek ψ_λ'yı bir anlamda birim vektör yapmaktır. Bu, şöyle anlatılır: Önce sabit bir $\lambda \in \mathbb{R}$ ve küçük bir $d\lambda > 0$ için (açıkça yazmasak da $d\lambda$'ya bağlı) bir $\Delta\Psi(q)$ fonksiyonu

$$\Delta\Psi(q) = \int_\lambda^{\lambda+d\lambda} \psi(t,q)dt \tag{6}$$

ile tanımlanır. Fizikçi $\psi_t = \psi(t,\cdot)$ fonksiyonları L^2'de olmasa da $\Delta\Psi$'nin L^2'de olduğunu gözler ve normalizasyon koşulunu (yani $c(\lambda)$ sabitinin seçimini)

$$\lim_{d\lambda\searrow 0} \frac{1}{d\lambda}\|\Delta\Psi\|^2 = \lim_{d\lambda\searrow 0} \frac{1}{d\lambda}\int_{\mathbb{R}} |\Delta\Psi(q)|^2 dq = 1 \tag{7}$$

olması koşuluyla belirler.

Burada kalkülüs sezgileriyle düşünenlerin kafasına şu soru takılabilir: $d\lambda$'ya bağlı olarak tanımladığımız $\Delta\Psi$ fonksiyonunun büyüklüğünün (6) ifadesine bakınca $d\lambda$ mertebesinde olması beklenir. O halde $\|\Delta\Psi\|^2$'nin de $(d\lambda)^2$ mertebesinde olacağı düşünülür ki bu durumda (7) olanaksızdır. Ne var ki farklı λ'lar için ψ_λ'ların birbirine "dik" olması aslında $\|\Delta\Psi\|^2 = O(d\lambda)$ olmasına neden olmaktadır ve bu da yukarıdaki prosedürün sorunsuzca işlemesini sağlar. Bunun fizikçi gözünden açıklaması için okuyucuyu Fok'un kitabına yönlendiriyoruz. Biz ileride matematiksel modeli tam olarak yerine oturtunca yukarıda olan-bitenin ne olduğu konusunda görüşümüz netleşecektir.

Fizikçinin yöntemini izleyip normalizasyonu yapalım: Aslında $\psi(t,q)$ içindeki $c(t)$ sabiti t'ye bağlı olmakla beraber bu bağımlılığın sürekli olduğunu varsayar ve zaten $d\lambda \to 0$ limitine geçeceğimizi gözlersek, tüm $t \in [\lambda, \lambda + d\lambda]$ için $c(t) = c(\lambda)$ alarak hesap yapabiliriz (hata terimi limitte yok olacaktır), dolayısıyla

$$
\begin{aligned}
\Delta\Psi(q) &= \int_{\lambda}^{\lambda+d\lambda} \psi(t,q)dt = c(\lambda)\int_{\lambda}^{\lambda+d\lambda} e^{itq/\hbar}dt \\
&= \frac{c(\lambda)\hbar}{iq}e^{itq/\hbar}\Big|_{t=\lambda}^{t=\lambda+d\lambda} \\
&= \frac{c(\lambda)\hbar}{iq}e^{i\lambda q/\hbar}\left[e^{i(d\lambda)q/\hbar}-1\right] \\
&= \frac{c(\lambda)\hbar}{iq}e^{i(\lambda+\frac{d\lambda}{2})q/\hbar}\left[e^{i\frac{d\lambda}{2}q/\hbar}-e^{-i\frac{d\lambda}{2}q/\hbar}\right] \\
&= \frac{2c(\lambda)\hbar}{q}e^{i(\lambda+\frac{d\lambda}{2})q/\hbar}\frac{e^{i\frac{d\lambda}{2}q/\hbar}-e^{-i\frac{d\lambda}{2}q/\hbar}}{2i} \\
&= \frac{2c(\lambda)\hbar}{q}e^{i(\lambda+\frac{d\lambda}{2})q/\hbar}\sin\left(\frac{d\lambda}{2\hbar}q\right)
\end{aligned}
$$

elde ederiz; burada $\sin z = \frac{e^{iz}-e^{-iz}}{2i}$ olduğunu kullandık. Demek ki, $c(\lambda)$ yerine kısaca c yazarak

$$
\begin{aligned}
\frac{1}{d\lambda}\int_{\mathbb{R}} |\Delta\Psi(q)|^2 dq = \frac{4c^2\hbar^2}{d\lambda}\int_{-\infty}^{\infty}\frac{\sin^2\left(\frac{d\lambda}{2\hbar}q\right)}{q^2}dq \\
\overset{u=\frac{d\lambda}{2\hbar}q}{=} 2\hbar c^2\int_{-\infty}^{\infty}\frac{\sin^2 u}{u^2}du = 2\pi\hbar c^2
\end{aligned}
$$

elde ederiz. Buradaki $\int_{-\infty}^{\infty}\frac{\sin^2 u}{u^2}du = \pi$ özdeşliği klasik bir kalkülüs alıştırmasıdır. Sonuç olarak

$$c = c(\lambda) = \frac{1}{\sqrt{2\pi\hbar}}$$

alarak istediğimiz normalizasyonu elde ediyoruz ve böylece fizikçi gözüyle

$$\{\psi(\lambda,q)\}_{\lambda\in\mathbb{R}} = \left\{\frac{1}{\sqrt{2\pi\hbar}}e^{i\lambda q/\hbar}\right\}_{\lambda\in\mathbb{R}}$$

topluluğu $L^2(\mathbb{R})$'nin P-özvektörlerinden oluşan bir ortonormal tabanı olmuş oluyor (bu noktada, 3 boyutta çalışıldığında $c(\lambda) = 1/(2\pi\hbar)^{3/2}$ çıktığını araya ekleyelim). Elde ortonormal taban olunca da, herhangi bir $\psi(q)$ dalga fonksiyonunun (5) türünden bir açılımı olması beklenir; yani

$$
\begin{aligned}
\psi(q) &= \int_{-\infty}^{\infty}\langle\psi,\psi_\lambda\rangle\,\psi_\lambda d\lambda \\
&=: \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\varphi(\lambda)e^{i\lambda q/\hbar}d\lambda \qquad (8)
\end{aligned}
$$

yazabiliriz, burada

$$\varphi(\lambda) = \langle\psi,\psi_\lambda\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\psi(q)e^{-i\lambda q/\hbar}dq \qquad (9)$$

ile verilen katsayılardır. Aslında φ fonksiyonu, matematikçilerin $\hat{\psi}$ ile göstermeye alışkın oldukları ve ψ'nin *Fourier dönüşümü* dedikleri fonksiyondur. Yukarıdaki (8) ise bizi $\varphi = \hat{\psi}$'den ψ'ye geri döndüren *ters Fourier dönüşümü*dür. Böylelikle Fourier dönüşümünün fizikçi sezgilerine dayanan geometrik bir yorumunu görmüş oluyoruz. Her ne kadar (8) ve (9) integralleri $\psi \in L^2$ için yakınsak olmayabilse de, bu integraller L^2 içinde yoğun olan $L^2 \cap L^1$ altuzayında anlamlıdır ve buradan yola çıkılarak Fourier dönüşümü tüm L^2'de tanımlı bir (doğrusal) izometriye genişletilebilir. Bu izometri L^2 üzerinde üniter bir dönüşüm tanımlar. Fourier dönüşümünün izometri olması yani

$$\int_{-\infty}^{\infty}|\psi(q)|^2 dq = \int_{-\infty}^{\infty}|\varphi(\lambda)|^2 d\lambda$$

olması, aslında (4) Pisagor-Parseval özdeşliğinin sürekli formudur.

Klasik Analiz'in önemli bir dalı olan Fourier Analizi konusunda ilgili okuyucuyu bu konudaki kaynaklara yönlendiriyoruz.

Yukarıdaki anlatıda bulduğumuz $\psi(\lambda,q)$ fonksiyonları P momentum operatörünün özvektörleriydi. Burada fizikçilerin geleneksel gösterimine geçerek λ yerine p yazacağız, dolayısıyla

$$\{\psi_p(q)\}_{p\in\mathbb{R}} = \left\{\frac{1}{\sqrt{2\pi\hbar}}e^{ipq/\hbar}\right\}_{p\in\mathbb{R}}$$

bizim P-özvektörlerden oluşan ortonormal tabanımız olacaktır. Burada hep $P\psi_p = \lambda\psi_p$ olduğu için $\psi_p(q)$ dalga fonksiyonu, kesin olarak p momentumuna sahip bir parçacığın durumunu temsil ediyor olarak düşünülebilir. Genel bir $\psi(q)$ durumu ise bu "kesin" durumların $\varphi(p)$ katsayılarıyla biraraya gelmiş hali olarak düşünülebilir; yani değişik olasılıkları bünyesinde barından karışık bir durum. Buradaki $\varphi(p) = \hat{\psi}(p)$ Fourier dönüşümü $\psi(q)$'yu tek olarak belirler, yani

$$\psi(q) \longleftrightarrow \varphi(p) \qquad (10)$$

eşleşmesini yapabiliriz. İşte bu iki temsilden soldakine fizikçiler "konum uzayındaki temsil", sağdakine de "momentum uzayındaki temsil" derler.

Daha önce söylediğimiz üzere, sistemin durumu farklı fiziksel büyüklüklerin fonksiyonları olarak değişik (ama denk!) şekillerde temsil edilebilir. Üstteki tartışma bize konum temsilinden momentum temsiline geçişin Fourier dönüşümü ile yapıldığını göstermektedir.

Daha önce demiştik ki, eğer a büyüklüğünü A operatörü ile ölçüyorsak ve dalga fonksiyonu $\psi(a)$ olarak verilmişse, A'nin etkisi $\psi(a) \mapsto a\psi(a)$ şeklindedir. Bu da demek oluyor ki P momentum operatörünün $\varphi(p)$ üzerindeki etkisi $\varphi(p) \mapsto p\varphi(p)$ şeklinde olmalıdır. Bunun böyle olduğunu hesaplayarak görelim:

Sistemimizin durumu (10) temsilleriyle verilsin, yani

$$\psi(q) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\varphi(p)e^{ipq/\hbar}dp \qquad (11)$$

olsun. Peki $P\psi$'nin momentum temsili ne olur? Görelim:

$$\begin{aligned} P\psi(q) &= \frac{1}{\sqrt{2\pi\hbar}}(-i\hbar)\frac{d}{dq}\int_{-\infty}^{\infty}\varphi(p)e^{ipq/\hbar}dp \\ &= \frac{1}{\sqrt{2\pi\hbar}}(-i\hbar)\int_{-\infty}^{\infty}\varphi(p)\frac{d}{dq}\left(e^{ipq/\hbar}\right)dp \\ & \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}p\varphi(p)e^{ipq/\hbar}dp \end{aligned}$$

olur. Burada türevle integralin değiştirilebildiğini varsaydık. En azından yeterince hızlı sönen φ fonksiyonları için bu doğrudur ve fiziksel olarak anlamlı dalga fonksiyonları bu koşulları sağlar. Yukarıdaki ifade (11) ile karşılaştırılınca görüyoruz ki

$$\psi(q) \leftrightarrow \varphi(p) \quad \text{ise} \quad P\psi(q) \leftrightarrow p\varphi(p).$$

Demek ki P'nin dalga fonksiyonunun momentum temsili $\varphi(p)$ üzerindeki etkisi gerçekten de $\varphi(p) \mapsto p\varphi(p)$ şeklinde! Fourier Analizi'ne aşina olan okur bunun türevin Fourier dönüşümüyle ilgili bilindik kurala karşılık geldiğini hatırlayacaktır.

3 Yoğun altuzaylarda tanımlı doğrusal dönüşümler

Şimdi artık işin matematik cephesine dönüp,

$$T : D(T) \subseteq H \to H \tag{12}$$

şeklinde, yoğun bir $D(T)$ altuzayında tanımlı doğrusal dönüşümleri irdeliyoruz. "Yoğun altuzayda tanımlı doğrusal dönüşüm" için YATDD kısaltmasını kullanacağız.

Burada T'nin sürekli olduğu *varsayılmamaktadır.* Aslında bu durum gereksizdir; çünkü T yoğun bir altuzayda sürekli ise, orada düzgün sürekli olduğu için (Teorem 1.8) ve H tam olduğundan, tek biçimde sürekli bir doğrusal $\tilde{T} : H \to H$ operatörüne genişler. Dolayısıyla T yerine tüm uzayda tanımlı ve sürekli $\tilde{T}$'yi irdeleyebiliriz. Dolayısıyla YATDD'ler süreksiz dönüşümler için anlamlı bir ayrımdır. Daha önce bahsi geçen P ve Q operatörleri bu sınıftandır.

Bundan böyle aksi belirtilmedikçe nesnemiz (12) ile verilen bir T dönüşümü olacaktır. İlk tanım-teoremimizi verelim:

Teorem 3.1. *Bir $y \in H$ için*

$$\forall x \in D(T) \qquad \langle Tx, y\rangle = \langle x, y^*\rangle$$

koşulunu sağlayan bir y^* elemanı varsa bu eleman tektir ve T^*y ile gösterilir. Bu özelliğe sahip y elemanları H'nin bir $D(T^*)$ altuzayını oluşturur ve

$$\begin{aligned} T^* : D(T^*) \subseteq H &\to H \\ y &\mapsto y^* \end{aligned}$$

dönüşümü doğrusaldır. Bu dönüşüme T'nin Hilbert eşleniği denir.

Teoremin kanıtı oldukça kolaydır. Burada y^* elemanının (varsa) tekliği için $D(T)$'nin yoğun bir altuzay olmasına başvuruyoruz. Çünkü bir Hilbert uzayında iki y_1 ve y_2 elemanı birbirine eşittir ancak ve ancak yoğun bir altuzaydan gelen tüm x elemanları için $\langle x, y_1\rangle = \langle x, y_2\rangle$ ise.

Tanım 3.2. $T = T^*$ ise T özeşleniktir denir.

Teorem 3.3. *T simetrik bir dönüşümse, yani*

$$\forall x, y \in D(T) \qquad \langle Tx, y\rangle = \langle x, Ty\rangle$$

ise $T^* \supseteq T$, yani, T^* dönüşümü T'nin bir genişlemesidir.

Kanıt. Bir $y \in D(T)$ verilsin ve sabitlensin. Üstteki eşitlik bize $y \in D(T^*)$ ve $T^*y = Ty$ olduğunu söyler, yani $T = T^*|_{D(T)}$.

Artık bu yeni tanıma göre Q ve P dönüşümlerimizin özeşlenik olduklarını kanıtlayabiliriz. İşe önce Q ile başlıyoruz. Biz yine yalınlık açısından kanıtı tek boyutta versek de, daha yüksek boyutlu durumdaki kanıt benzerdir: □

Teorem 3.4. *Tanım altuzayı*

$$D(Q) = \{\psi(q) \in L^2(\mathbb{R}) : q\psi(q) \in L^2(\mathbb{R})\}$$

olarak verilen

$$\begin{aligned} Q : D(Q) \subseteq L^2(\mathbb{R}) &\to L^2(\mathbb{R}) \\ \psi(q) &\mapsto q\psi(q) \end{aligned}$$

dönüşümü özeşleniktir.

Kanıt. Bu dönüşümün simetrikliği aynen Örnek 1.18'de olduğu gibi gösterilir. Demek ki $Q^* \supseteq Q$. Şimdi $Q \supseteq Q^*$ olduğunu göstereceğiz. Bunun için bir $y \in D(Q^*)$ alalım. Biz, $y \in D(Q)$ ve $Q^*y = Qy$ olduğunu göstermeliyiz.

Şimdi $y^* = Q^*y$ olsun. Yani her $x \in D(Q)$ için

$$\begin{aligned} \langle Qx, y\rangle &= \int_{\mathbb{R}} qx(q)\overline{y(q)}dq = \int_{\mathbb{R}} x(q)\overline{qy(q)}dq \\ &= \int_{\mathbb{R}} x(q)\overline{y^*(q)}dq = \langle x, y^*\rangle\,, \end{aligned}$$

dolayısıyla her $x \in D(Q)$ için

$$\int_{\mathbb{R}} x(q)\overline{[qy(q) - y^*(q)]}dq = 0. \tag{*}$$

Herhangi bir kompakt $[a, b]$ aralığı için

$$\chi_{[a,b]}(q) \cdot [qy(q) - y^*(q)] \quad \text{ve} \quad q\chi_{[a,b]}(q) \cdot [qy(q) - y^*(q)]$$

fonksiyonlarının da L^2'de olduğuna dikkat edelim. Demek ki (*) ifadesinde $x(q) = \chi_{[a,b]}(q) \cdot [qy(q) - y^*(q)]$ alabiliriz ve bu bize

$$\int_a^b |qy(q) - y^*(q)|^2 dq = 0$$

verir. Yani $[a, b]$'deki hemen her q için $qy(q) - y^*(q) = 0$ olmalıdır. Bu her $[a, b]$ için doğru olduğundan demek ki L^2'de $qy(q) = y^*(q) \in L^2$. Bu hem $y \in D(Q)$ olduğunu, hem de $Qy = y^* = Q^*y$ olduğunu kanıtlar ve işimiz biter. □

Sırada momentum operatörü P var. Aslında burada bir kestirme yol var: $\mathcal{F} : L^2 \to L^2$ ile Fourier dönüşümünü gösterirsek, bunun üniter oluşunu ve $P = \mathcal{F}^{-1}Q\mathcal{F}$ bağıntısını kullanarak aslında P dönüşümünün Q'nun bir "kılık değiştirmiş" hali olduğu, buna dayanarak da özeşlenik olması gerektiği ve ileride göreceğimiz spektral özelliklerinin de Q'nunkiyle birebir aynı olduğu görülebilir. Ancak hem bu geçiş için kullanılan Fourier dönüşümünün özelliklerini kanıtlamadığımız, hem de matematiksel açıdan bazı önemli kavramları sunmamıza aracı olacağı için doğrudan bir kanıtı da vermek istiyoruz.

Teorem 3.5. *Tanım uzayı*

$$D(P) = \{\psi(q) \in L^2(\mathbb{R}) : \ \psi \textit{ her kompakt aralıkta mutlak süreklidir ve } \psi'(q) \in L^2(\mathbb{R})\textit{'dir}\}$$

olarak verilen $P = -i\hbar\frac{d}{dq}$ *dönüşümü özeşleniktir.*

Kanıta geçmeden önce açıklamak istediğimiz birkaç nokta var:

Öncelikle, L^2'nin elemanları denklik sınıfları olduğuna göre, türev alma ile ne kastedilmektedir? Örneğin 0 sabit fonksiyonu ile $\chi_{\mathbb{Q}}$ fonksiyonu $L^2(\mathbb{R})$'de 0 elemanının denklik sınıfındadırlar. İlki her yerde sonsuz kez türevliyken ikincisi tek bir noktada dahi sürekli değildir. Peki türev işlemini bunlardan hangisi üzerinde tanımlayacağız?

Burada şunu gözlüyoruz: L^2'deki her bir denklik sınıfında en fazla bir tane sürekli (ve dolayısıyla en fazla bir tane türevli) fonksiyon olabilir. İşte bir L^2 elemanının mutlak sürekli, türevli vs. olduğu söylenirken veya böyle bir elemanın türevi alınırken, bu biricik temsilciye atıfla konuşulmaktadır. Örneğin $\chi_{[0,\infty)}$ elemanının denklik sınıfında sürekli fonksiyon yoktur, bu nedenle bu fonksiyon (denklik sınıfı) L^2'nin süreksiz bir elemanıdır diyebiliriz.

Üstte geçen bir diğer kavram da mutlak sürekliliktir. Mutlak süreklilik düzgün sürekliliği gerektirir ama bundan çok daha fazlasıdır. Mutlak sürekli fonksiyonlar, (Lebesgue) integrallenebilir bir fonksiyonun integrali olarak yazılabilen fonksiyonlardır; eşdeğer tanımla hemen her yerde türevlenebilir olup kendi türevlerinin integrali olan fonksiyonlardır. Mutlak sürekli bir x fonksiyonu

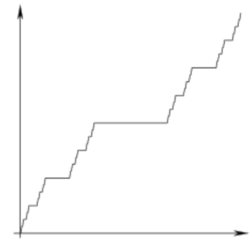
$$x(q) = x(q_0) + \int_{q_0}^{q} x'(t)dt$$

diye yazabileceğimiz Kalkülüs'ün Temel Teoremi'nin sağlar. Bu fonksiyonların çarpan olduğu ifadelerin integrallerinde kısmi integrasyon (integration by parts) yöntemi uygulanabilir.

Akla "hemen herde türevi olup da kendi türevinin integrali olmayan fonksiyonlar var mıdır?" sorusu gelebilir. Örnek olarak, $[0,1]$ aralığından kendisine giden sürekli bir fonksiyon olan "Şeytan Merdiveni" (Devil's Staircase) fonksiyonunu verebiliriz. Bu fonksiyon Cantor kümesinden $[0,1]$ aralığına, Cantor kümesi noktalarının 3'lü tabanda sadece 0 ve 2'lerle tek olarak yazılabilmesi gözleminden yola çıkılarak

$$(0.a_1a_2a_3\cdots)_3 \mapsto (0.\frac{a_1}{2}\frac{a_2}{2}\frac{a_3}{2}\cdots)_2$$

şeklinde tanımlanan fonksiyonun, Cantor kümesinin dışındaki (silinen) aralıklarda sabit olarak tanımlanarak sürekli şekilde $[0,1]$'e genişletilmesiyle elde edilir (bkz. alttaki resim).



Silinen açık aralıkların toplam uzunluğu 1 olduğu için bu fonksiyon $[0,1]$ aralığında hemen her yerde türevlidir ve türevi 0'dır; ama elbette sabit fonksiyon olmadığı için kendi türevinin integrali değildir.

Teorem 3.5'nin kanıtından önce son bir notu da şöyle düşelim: Daha yüksek boyutta $D(P)$'nin tanımında ψ'nin her argümanına göre kompakt aralıklarda ayrı ayrı mutlak sürekli olması ve kısmi türevlerin L^2'de olması istenir.

Kanıttan önce bir lemmamız olacak:

Lemma 3.6. $\psi \in D(P)$ *ise* $\lim_{|q|\to\infty} \psi(q) = 0$.

Kanıt. Bu sonucun zaten fiziksel olarak anlamlı sistemlerin dalga fonksiyonları için beklenen bir özellik olduğunu söyleyerek başlayalım. Biz $q \to +\infty$ limiti için bir kanıt vereceğiz, diğer durum benzerdir:

Herhangi bir $a > 0$ için $\psi\chi_{[0,a]} \in L^2(\mathbb{R})$ olur, böylece

$$\begin{aligned}
2Re\left\langle \psi\chi_{[0,a]}, \psi' \right\rangle &= \left\langle \psi\chi_{[0,a]}, \psi' \right\rangle + \left\langle \psi', \psi\chi_{[0,a]} \right\rangle \\
&= \int_0^a (\psi(q)\overline{\psi'(q)} + \psi'(q)\overline{\psi(q)})dq \\
&= \frac{1}{2}\int_0^a \frac{d}{dq}\left(|\psi(q)|^2\right) dq \\
&= \frac{1}{2}(|\psi(a)|^2 - |\psi(0)|^2)
\end{aligned}$$

gelir. Şimdi $a \to +\infty$ iken $\psi\chi_{[0,a]} \xrightarrow{L^2} \psi$ olduğundan üstte en baştaki terimde limit vardır, o halde en son satırda da limit olmalıdır. Bu da $\lim_{a\to+\infty} |\psi(a)|$ limitinin var olduğunu gösterir ama $\psi \in L^2$ olduğundan bu limit 0 olmalıdır. $\square$

Teorem 3.5'in kanıtı. Önce P'nin simetrik olduğunu görelim: Herhangi iki $x, y \in D(P)$ için

$$\begin{aligned}
\langle Px, y \rangle &= -i\hbar \int_{-\infty}^{\infty} x'(q)\overline{y(q)}dq \\
&= -i\hbar x(q)\overline{y(q)}\Big|_{-\infty}^{\infty} + i\hbar \int_{-\infty}^{\infty} x(q)\overline{y'(q)}dq \\
&\overset{\text{Lemma}}{=} (0-0) + \int_{-\infty}^{\infty} x(q)\overline{(-i\hbar)y'(q)}dq \\
&= \langle x, Py \rangle .
\end{aligned}$$

Demek ki $P^* \supseteq P$. Şimdi tersini görelim: Bir $y \in D(P^*)$ alalım. Amacımız $y \in D(P)$ ve $Py = P^*y$ olduğunu göstermek.

Şimdi $P^*y =: y^*$ diyelim. Herhangi bir $[a,b]$ kompakt aralığı seçip sabitleyelim ve $D(P)$'den desteği bu aralıkta kalan bir x fonksiyonu alalım. Elbette x' fonksiyonunun da desteği bu aralıkta kalacaktır. O halde böyle bir x için

$$\langle Px, y\rangle = -i\hbar \int_a^b x'(q)\overline{y(q)}dq = \langle x, y^*\rangle$$
$$= \int_a^b x(q)\overline{y^*(q)}dq \qquad (*)$$

olur. Şimdi bu aralıkta türevi $\overline{y^*}$ olan bir z fonksiyonu tanımlayacağız. Her $q \in [a,b]$ için

$$z(q) := c + \int_a^q \overline{y^*(t)}dt$$

ile tanımlansın. Buradaki c sabitinin nasıl seçileceği ileride açıklanacaktır. L^2 fonksiyonlarının kompakt aralıklara kısıtlanışları (Hölder Eşitsizliği'ne göre) integrallenebilirdir; dolayısıyla üstteki integral anlamlıdır. Bu durumda z fonksiyonu $[a,b]$ üzerinde mutlak sürekli ve $z' = \overline{y^*}$ olur. Böylece (*)'da kısmi integrasyon ile devam ederek

$$\langle Px, y\rangle = -i\hbar \int_a^b x'(q)\overline{y(q)}dq = \int_a^b x(q)\overline{y^*(q)}dq$$
$$= x(q)z(q)\Big|_a^b - \int_a^b x'(q)z(q)dq = -\int_a^b x'(q)z(q)dq$$

gelir. Toparlarsak, desteği $[a,b]$'de kalan her $x \in D(P)$ için

$$\int_a^b x'(q)[z(q) + \overline{i\hbar y(q)}]dq = 0 \qquad (**)$$

olur. Amacımız, desteği $[a,b]$'de olan ve bu aralıkta $x'(q) = \overline{[z(q) + \overline{i\hbar y(q)}]}$ şartını sağlayan bir $x \in D(P)$ bulmak. Elbette bu x için adayımız

$$x(q) = \begin{cases} \int_a^q \overline{[z(t) + \overline{i\hbar y(t)}]}dt, & q \in [a,b] \\ 0, & q \notin [a,b] \end{cases}$$

olacaktır. Öncelikle bu tanımdaki integralin anlamlı olduğuna ve $x \in D(P)$ koşulunun sağlandığına dikkat edelim. Elbette $q \le a$ için $x(q) \equiv 0$. Biz bunun $q \ge b$ için de olmasını istiyoruz. İşte z'nin tanımındaki c sabitini tam da üstteki tanımda $x(b) = 0$ olacak şekilde seçelim. Bu durumda $x \in D(P)$ ve desteği $[a,b]$'de kalmış olur; üstelik $[a,b]$'de $x'(q) = \overline{[z(q) + \overline{i\hbar y(q)}]}$ olduğundan (**) ile

$$\int_a^b |z(q) + \overline{i\hbar y(q)}|^2 dq = 0$$

gelir. Bu da $[a,b]$'de $z + \overline{i\hbar y} = 0$ olduğu anlamına gelir. Demek ki y fonksiyonu $[a,b]$'de mutlak süreklidir ve türev alarak $[a,b]$'de

$$z' = \overline{y^*} = -\overline{i\hbar y}$$

yani $y^* = -i\hbar y'$ elde ederiz. Seçtiğimiz $[a,b]$ aralığı rastgele olduğundan bu tüm $\mathbb{R}$'de geçerlidir. Demek ki hem y tüm kompakt aralıklarda mutlak süreklidir hem de türevi L^2'dedir. Sonuçta $y \in D(P)$ ve üstelik $P^*y = y^* = -i\hbar y' = Py$. Bu da kanıtı bitirir. □

4 Spektrum

Bu kısımda X yine bir kompleks Banach uzayı ve $T : D(T) \subseteq X \to X$ bir YATDD olsun. Verilen bir $\lambda \in \mathbb{C}$ ve $y \in X$ için

$$(T - \lambda)x = y$$

denkleminin çözülebilirliği problemini irdeliyoruz. Olabilecek en iyi senaryo bu denklemin tüm y'ler için biricik bir çözümü olması ve bu çözümü veren sürekli bir $(T-\lambda)^{-1}$ ters dönüşümü olmasıdır.

Ne yazık ki her λ için işler bu kadar yolunda gitmeyebilir. Sonuç olarak şu 4 durumdan biri ve yalnız biri geçerli olacaktır:

1. $T - \lambda$ birebir değildir, yani $(T - \lambda)x = 0$ olan bir $x \neq 0$ vardır. Bu durumda ters dönüşüm söz konusu olamaz. Bunun olduğu λ değerleri tam da T'nin özdeğerleridir. Bu durumda $\lambda \in \sigma_p(T)$ deriz ("point spectrum").

2. $T - \lambda$ birebirdir ama görüntü uzayı $R(T - \lambda)$ "çok küçüktür", yani X'te yoğun bir altuzay değildir. Bu da denklemin yok denecek kadar az bir y'ler kümesi için çözülebilir olması demektir. Bu durumda $\lambda \in \sigma_r(T)$ diyeceğiz ("residual spectrum").

3. $T - \lambda$ birebirdir ve $R(T - \lambda)$ görüntüsü yoğun bir altuzaydır; ama $(T - \lambda)^{-1} : R(T - \lambda) \to X$ sürekli değildir. Bu durumda $\lambda \in \sigma_c(T)$ denir ("sürekli/continuous spectrum").

4. Yukarıdaki üç durum dışındaki hallerde $\lambda \in \rho(T)$ ("resolvent set") diyoruz. Bu, denklemin hemen hemen her y için, y'nin sürekli bir fonksiyonu ile çözülebildiği anlamına gelir.

Tanım 4.1. $\sigma(T) := \sigma_p(T) \sqcup \sigma_r(T) \sqcup \sigma_c(T)$ kümesine T'nin spektrumu denir.

Devam etmeden önce önemli bir kavram olan kapalı dönüşümleri tanıtalım:

Tanım 4.2. X ve Y norm uzayları ve $T : D(T) \subseteq X \to Y$ doğrusal olsun. Eğer T'nın grafı, yani

$$\Gamma_T = \{(x, Tx) : x \in D(T)\}$$

kümesi $X \times Y$ içinde kapalı bir altuzaysa T'ye kapalı bir dönüşümdür denir.

Bunun (topolojiden çağrışım yapacağı üzere) kapalı kümeleri kapalı kümelere göndermekle bir ilgisi yoktur. Kapalı dönüşümler sürekli olabildikleri gibi, olmasalar bile onları kullanışlı kılan bazı "iyi" özelliklere sahiptirler.

Teorem 4.3. *$T : D(T) \subseteq X \to X$ doğrusal dönüşümü için (a) T'nin kapalı, veya, (b) $D(T) = X$ ve T'nin sürekli olduğunu varsayalım. Her iki durumda da $\lambda \in \rho(T)$ için $(T - \lambda)^{-1}$ tüm X'te tanımlıdır.*

Kanıt. Tüm uzayda tanımlı sürekli dönüşümler kapalı olacağı için (b) durumu (a)'dan hemen çıkar. (a)'nın kanıtı içinse şunları gözlemek yeterlidir: T kapalı olduğu için $T-\lambda$ da kapalıdır. Kapalı ve birebir dönüşümlerin tersi de kapalıdır (ters dönüşümün grafı düz dönüşümünkinin bir ayna yansıması gibidir), dolayısıyla $(T-\lambda)^{-1} : R(T-\lambda) \subseteq X \to X$ kapalıdır. Ama bu dönüşüm bir yandan da sürekli olduğu için, kapalılık, tanım altuzayının kapalı altuzay olmasını gerektirir. Ama $R(T-\lambda)$ yoğun bir altuzay olduğu için kapalı olması $R(T-\lambda) = X$ demektir. □

Teorem 4.4. *Bir önceki teoremdeki her iki durumda da $\rho(T)$ açık bir kümedir. Üstelik (b) durumunda*

$$\emptyset \neq \sigma(T) \subseteq \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}.$$

Kanıt. Her iki senaryoda da $\lambda \in \rho(T)$ için $(T-\lambda)^{-1}$ tüm X'te tanımlı ve süreklidir. Şimdi $\lambda \in \rho(T)$ ve $\mu \in \mathbb{C}$ için

$$T - \mu = T - \lambda + \lambda - \mu = (T-\lambda)(1 - (\mu - \lambda)(T-\lambda)^{-1})$$

yazabiliriz. Şimdi (1) özdeşliğini hatırlarsak görürüz ki eğer

$$|\mu - \lambda| \|(T-\lambda)^{-1}\| < 1$$

ise $(1 - (\mu-\lambda)(T-\lambda)^{-1})$ dönüşümünün ve dolayısıyla $T-\mu$'nün X'te tanımlı sürekli bir tersi vardır. O halde $\lambda \in \rho(T)$ iken λ'ya yeterince yakın μ sayıları da aynı kümededir; o halde $\rho(T)$ açıktır.

T sürekli iken $|\lambda| > \|T\|$ için $\lambda \in \rho(T)$ olması esasen (2) özdeşliğinden gelmektedir. Dolayısıyla tüm uzayda tanımlı sürekli operatörlerin spektrumu kompakttır. Spektrumun boş olmamasının kanıtı şöyle özetlenebilir: Eğer boş olsaydı $\rho(T) = \mathbb{C}$ olurdu. Ama herhangi bir $\lambda \in \mathbb{C} = \rho(T)$ civarındaki μ değerleri için $(T-\mu)^{-1}$ ters dönüşümü T'nin bir kuvvet serisiyle verilebilir (hemen üstteki hesapları gözden geçiriniz). Bu da herhangi bir $x \in X$ ve sürekli doğrusal $f : X \to \mathbb{C}$ için $\mu \mapsto f((T-\mu)^{-1}x)$ dönüşümünü tüm $\mathbb{C}$'de holomorf yapar. Ancak $|\mu| \to \infty$ iken $\|(T-\mu)^{-1}\| \to 0$ olduğundan bu holomorf fonksiyon sınırlıdır, dolayısıyla Liouville Teoremi'ne göre sabit olmalıdır (ki bu sabit değer de mecburen 0'dır). Bunun her x ve f için doğru olması bir çelişkidir. □

Daha önce kuvvet serileri yardımıyla sürekli bir $T : X \to X$ için $f(T)$ dönüşümünün nasıl tanımlanabileceğini görmüştük. Spektrumu tanımladıktan sonra yine Kompleks Analiz'e dayanan bir yol daha elimizde olacak:

$T : X \to X$ sürekli iken spektrumunun kompakt olacağını gözlemiştik. Şimdi $U \supseteq \sigma(T)$ bir açık küme ve $f : U \to \mathbb{C}$ holomorf olsun. Şimdi U içinde $\sigma(T)$ kümesini pozitif yönde çeviren bir γ basit kapalı eğrisi alalım. Hatırlanırsa bu eğri içinde kalan bir z_0 noktası için Cauchy integral formülü bize

$$f(z_0) = \frac{1}{2\pi i} \oint_\gamma \frac{f(z)}{z - z_0} dz$$

olduğunu söyler. Bu integralin yakınsamasına dair yapılan kanıt aynen kuvvet serilerinde olduğu gibi operatörlere taşınıp

$$f(T) := \frac{1}{2\pi i} \oint_\gamma \frac{f(\lambda)}{\lambda - T} d\lambda$$

integralinin sürekli bir $f(T)$ dönüşümüne yakınsadığı gösterilebilir (bu integral bir Riemann integralidir, tek fark, kısmi toplamlar sayı değil operatör değerlidir ve operatörler uzayında bir limite yakınsar). Burada elbette $\frac{1}{\lambda - T} := (\lambda - T)^{-1}$ olarak anlaşılacaktır. Eğrimiz spektruma değmediği için eğri boyunca bu ifade anlamlıdır. Üstelik, γ eğrisi söylenen şartları sağladığı sürece integralin sonucu eğrinin seçiminden bağımsızdır (tıpkı Cauchy integralinde olduğu gibi) çünkü spektrumdan hiçbir nokta içermeyen bir bölgeyi çeviren bir eğride iç bölgede $f(\lambda)/(\lambda - T)$ holomorf olacak ve integral 0 çıkacaktır. Dolayısıyla γ'yı spektruma değdirmeden deforme etmek integralin değerini değiştirmez.

İleride göreceğiz ki, özeşlenik dönüşümler için bunun çok daha fazlasını yapabiliyoruz: Sadece holomorf değil, $\mathbb{R}$'de sürekli (hatta sadece Lebesgue ölçülebilir ve sınırlı) f fonksiyonları için dahi $f(T)$'yi tanımlamak mümkün olacak. Özeşlenik dönüşümlere odaklanmadan önce Spektral Dönüşüm Teoremi'ni de (Spectral Mapping Theorem) söyleyelim:

Teorem 4.5 (Spektral Dönüşüm Teoremi)**.** *Sürekli doğrusal bir $T : X \to X$ dönüşümü ve $\sigma(T)$'yi içeren açık bir U kümesinde holomorf bir $f : U \to \mathbb{C}$ için*

$$\sigma(f(T)) = f(\sigma(T)). \tag{*}$$

Bu teoremin polinom $f(\lambda)$ fonksiyonları için kanıtı, Cauchy integrallerine girmeden de verilebilir:

$$f(T) - f(\lambda) = (T - \lambda) g(T, \lambda) = g(T, \lambda)(T - \lambda)$$

yazabiliriz; burada $g(T, \lambda)$ hem T hem λ'nın bir polinomudur. Sol tarafın $X \to X$ birebir, örten olup sürekli bir tersinin olmasının ancak ve ancak sağdaki $T - \lambda$ çarpanı için aynısının doğru olmasıyla mümkün olduğu görülebilir, yani $f(\lambda) \in \rho(f(T))$ ancak ve ancak $\lambda \in \rho(T)$. Bu da (*)'a denktir. Genel durumda ise Cauchy integral temsili üzerinden benzer bir çarpanlara ayırma işlemi uygulanarak sonuca gidilir.

4.1 Özeşlenik dönüşümlerde spektrum, Q ve P örnekleri

Bu bölümde yine $T : D(T) \subseteq H \to H$ şeklinde bir YATDD üzerine konuşacağız.

Teorem 4.6. *T^* kapalı bir dönüşümdür.*

Kanıt. Bize T^*'in grafından $(y_n, T^* y_n)$ şeklinde bir dizi verilsin ve bu dizi $H \times H$'de bir (y, z) elemanına yakınsasın. Bu durumda (y, z)'nin de grafta olduğunu yani $y \in D(T^*)$ ve $z = T^* y$ olduğunu göstermemiz gerek. Dikkat edilirse varsayımımıza göre $y_n \to y$ ve $T^* y_n \to z$. Şimdi herhangi bir $x \in D(T)$ için

$$\langle Tx, y \rangle = \lim_n \langle Tx, y_n \rangle \overset{y_n \in D(T^*)}{=} \lim_n \langle x, T^* y_n \rangle = \langle x, z \rangle .$$

Demek ki $y \in D(T^*)$ ve $T^* y = z$. □

Doğal Sonuç 4.7. *T özeşlenik ise kapalı bir dönüşümdür. Dolayısıyla $\sigma(T)$ kapalı bir kümedir ve her $\lambda \in \rho(T)$ için $(T-\lambda)^{-1}$ tüm H'de tanımlı ve süreklidir.*

Daha önce özeşlenik dönüşümlerin özdeğerlerinin gerçel olması gerektiğini görmüştük. Aslında sadece özdeğerler değil, tüm spektrum $\mathbb{R}$'de kalmalıdır:

Teorem 4.8. *$T : D(T) \subseteq H \to H$ özeşlenik ise*
(i) $\sigma(T) \subseteq \mathbb{R}$,
(ii) $\sigma_r(T) = \emptyset$,

Kanıt. (i) $\lambda = \alpha + i\beta$ alalım, burada $\alpha, \beta \in \mathbb{R}$ ve $\beta \neq 0$ olsun. Şimdi

$$\begin{aligned}\|(T-\lambda)x\|^2 &= \langle (T-\lambda)x, (T-\lambda)x\rangle \\ &= \langle T-\alpha x - i\beta x, T - \alpha x - i\beta x\rangle\end{aligned}$$

ifadesini açarak, çeşitli sadeleştirmelerden sonra

$$\|(T-\lambda)x\|^2 \geq |\beta|^2\|x\|^2$$

elde ederiz. Ama $\beta \neq 0$ olduğundan bu $T-\lambda$'nın birebir olduğu ve tersinin de sürekli olduğu anlamına gelir. Ancak $R(T-\lambda)$ yoğun bir altuzay olmalıdır; eğer olmasaydı bir $0 \neq z \perp R(T-\lambda)$ bulunabilirdi. O zamanda her $x \in D(T)$ için $0 = \langle (T-\lambda)x, z\rangle$, yani

$$\langle Tx, z\rangle = \langle \lambda x, z\rangle = \langle x, \overline{\lambda} z\rangle$$

olurdu. Bu da $z \in D(T^*)$ ve $T^* z = \overline{\lambda} z$ demek olur. Ama $T = T^*$ olduğundan bunun anlamı $\overline{\lambda}$'nın bir T-özdeğer olmasıdır. Halbuki $\overline{\lambda} \notin \mathbb{R}$, çelişki. O halde $R(T-\lambda)$ yoğundur ve $(T-\lambda)^{-1}$ süreklidir, bu da $\lambda \in \rho(T)$ demektir.

(ii) $\lambda \in \sigma(T) \backslash \sigma_p(T)$ olsun. Artık $\lambda \in \mathbb{R}$ olduğunu biliyoruz. Bu durumda $T-\lambda$ birebirdir. Yine $R(T-\lambda)$'nın yoğun olması gerektiğini iddia ediyoruz, bu aynen (i)'de olduğu gibi kanıtlanır (eğer yoğun olmasaydı $\overline{\lambda} = \lambda$ özdeğer olurdu). Ama o zaman da $\lambda \in \sigma_r(T)$ olamaz, $\lambda \in \sigma_c(T)$ olmalıdır. □

Teorem 4.9. *Kuantum Mekaniği'nin Q konum ve P momentum operatörler için*

$$\sigma(Q) = \sigma_c(Q) = \sigma(P) = \sigma_c(P) = \mathbb{R}.$$

Kanıtı önce Q için verelim. Aslında Q için problem çözülünce, aradaki Fourier bağlantısından P'yi de çözmüş oluruz; ama biz yine de doğrudan bir kanıt vereceğiz.

Biz zaten Q'nun özeşlenik olduğunu ve özdeğerinin olmadığını biliyoruz. Artık spektrum (residual spectrum) da boş olduğuna göre (Teorem 4.8) demek ki spektrum sadece sürekli spektrumdan oluşur ve bu da $\mathbb{R}$'nin içinde bir kapalı küme olmak zorundadır. Biz her $\lambda \in \mathbb{R}$'nin spektrumda olduğunu iddia ediyoruz: Bir λ seçelim ve $(Q-\lambda)^{-1}$ dönüşümüne bakalım. Bunun süreksiz olduğunu iddia ediyoruz. Elbette ters dönüşüm, tanımlı olduğu y'lerde

$$(Q-\lambda)^{-1} y(q) = \frac{1}{q-\lambda} y(q)$$

ile veriliyor olmalıdır. Biz, desteği $\frac{1}{q-\lambda}$'nın patlama noktası olan λ'ya yakın y fonksiyonları ile çalışırsak, ters dönüşümün bu vektörleri çok daha büyüterek resmetmesini kullanarak süreksizliği gösterebiliriz. Bu fikirden ilham alarak her $\varepsilon > 0$ için $y_\varepsilon = \chi_{[\lambda+\varepsilon, \lambda+\varepsilon+1]}$ alalım. Bu durumda

$$x_\varepsilon(q) := \frac{1}{q-\lambda} y_\varepsilon(q) = \begin{cases} \frac{1}{q-\lambda}, & q \in [\lambda+\varepsilon, \lambda+\varepsilon+1] \\ 0, & d.d. \end{cases}$$

fonksiyonu desteği $[\lambda+\varepsilon, \lambda+\varepsilon+1]$ aralığında kalan sınırlı bir fonksiyondur. Buradan $x_\varepsilon \in D(Q)$ olduğu kolaylıkla görülebilir; üstelik $(Q-\lambda)x_\varepsilon = y_\varepsilon$. Ama yine kolayca görüleceği üzere

$$\|y_\varepsilon\|^2 = \int_{\lambda+\varepsilon}^{\lambda+\varepsilon+1} dq = 1$$

ama

$$\begin{aligned}\|x_\varepsilon\|^2 = \|(Q-\lambda)^{-1} y_\varepsilon\|^2 &= \int_{\lambda+\varepsilon}^{\lambda+\varepsilon+1} \frac{1}{(q-\lambda)^2} dq \\ &= \int_{\varepsilon}^{\varepsilon+1} \frac{1}{u^2} du \xrightarrow{\varepsilon \searrow 0} +\infty.\end{aligned}$$

Sonuçta $\|(Q-\lambda)^{-1} y_\varepsilon\| / \|y_\varepsilon\|$ oranı üstten sınırsız olduğuna göre $(Q-\lambda)^{-1}$ süreksizdir ve $\lambda \in \sigma_c(Q)$ olur.

Momentum dönüşümü P içinse, başta Q için söylenenlerin aynısı geçerlidir. Yani iş sadece her gerçel λ sayısının sürekli spektrumda olduğunu göstermeye kalır. Bir $\lambda \in \mathbb{R}$ verilsin ve $n \in \mathbb{N}$ için

$$x_n(q) = \begin{cases} e^{i\lambda q}, & |q| \leq n \\ 0, & |q| \geq n+1 \end{cases}$$

olarak tanımlayalım ve geriye kalan $[-n-1, -n]$ ile $[n, n+1]$ aralıklarında x_n'i iki uç arasında doğrusal biçimde bağlayarak sürekli bir fonksiyon elde edelim. Bu durumda $x_n \in D(P)$ olacaktır. Dikkat edilirse $|q| \leq n$ için $|x_n(q)| = 1$, bu da

$$\|x_n\| = \sqrt{\int_{\mathbb{R}} |x_n(q)|^2 dq} \geq \sqrt{\int_{-n}^{n} dq} \geq \sqrt{2n}$$

alt sınırını verir. Öte yandan $(P-\lambda)x_n$ fonksiyonu $|q| < n$ ve $|q| > n+1$ için sıfırdır, arada kalan, uzunluğu 1 olan iki aralıkta da hem λx_n hem x_n' türevi, n'den bağımsız mutlak sabitlerle sınırlı olduğundan bir mutlak $C > 0$ için $\|(P-\lambda)x_n\| \leq C$ olduğu gösterilebilir. Dolayısıyla $y_n = (P-\lambda)x_n$ alınarak $\|(P-\lambda)^{-1} y_n\| / \|y_n\| \geq \sqrt{2n}/C$ oranının üstten sınırsızlığı ile $(P-\lambda)^{-1}$ dönüşümünün süreksizliği kanıtlanır. Dolayısıyla $\lambda \in \sigma_c(P)$.

5 Özeşlenik dönüşümler için spektral temsil teoremi

Bu kısımda özeşlenik dönüşümleri birer Riemann-Stieltjes integrali ile temsil etmemizi sağlayan von Neumann'ın ünlü teoremine değineceğiz. İlk alt kısımda Riemann-Stieltjes integrali hakkında bir özet geçiyoruz.

5.1 Riemann-Stieltjes integrali

Önce en basit haliyle Riemann integralinin temel fikrini hatırlayalım: Elimizde bir kompakt $[a, b]$ aralığı ve sürekli bir $f : [a, b] \to \mathbb{R}$ fonksiyonu var. Bu aralığın bir

$$a = x_0 < x_1 < \cdots < x_n = b$$

parçalanışını alıyoruz ve bu parçalanışa karşılık, adına Riemann toplamı dediğimiz

$$\sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1})$$

toplamını oluşturuyoruz. Amacımız, parçalanış inceldikçe bu toplama ne olduğuna bakmak. Burada $f(x_{i-1})$ teriminde x_{i-1} yerine $[x_{i-1}, x_i]$ aralığından herhangi bir nokta da seçilebilirdi; sonuçta aralıklar küçüldükçe f'nin $[a, b]$'deki düzgün sürekliliği sayesinde toplama yansıyan fark limitte sıfıra gidecektir.

Sonuçta gösterebiliyoruz ki parçalanış inceldikçe (bunu doğru olmasa da $\lim_{n\to\infty}$ olarak yazacağız, kastedilen parçalanışın incelmesidir) bu toplamlar belli bir limit değere gider ve bu limit değeri

$$\int_a^b f(x)dx := \lim_{n\to\infty} \sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1})$$

Riemann integrali olarak tanımlarız. Buradaki $f(x_{i-1})(x_i - x_{i-1})$ terimini, f fonksiyonunun $[x_{i-1}, x_i]$ aralığındaki değeri ile o aralığın "büyüklüğü"nün çarpımı olarak okuyabiliriz. Bu bağlamda büyüklük demek uzunluk demektir. Yani Riemann toplamı (ve integrali), f'nin aldığı değerlerin bir tür ağırlıklı toplamıdır. Ölçü Kuramı dilinde kümelerin "büyüklüğünü" ya da "ağırlığını" ölçen aracın ölçüdür. Bir gerçel sayı aralığına büyüklük olarak o aralığın uzunluğunu biçen ölçüye Lebesgue ölçüsü denir. Yani Ölçü Kuramı söylemiyle biz yukarıda f'nin Lebesgue ölçüsüne göre integralini almış olduk.

Şimdi aklımıza yeni bir fikir getirelim: Bir aralığın büyüklüğünü onun uzunluğuyla değil de başka şekilde ölçen bir araç (ölçü) olsa ne olur? Aslında istenirse aralığa negatif bir ağırlık bile atamayı düşünebiliriz ama biz basit düşünelim ve aralıkların ağırlığının/büyüklüğünün negatif olmamasını isteyelim. Böyle bir ölçüm mekanizmasının akla gelen ilk yollarından biri şudur: Azalmayan (non-decreasing) bir $g : \mathbb{R} \to \mathbb{R}$ fonksiyonu alınır ve $(x, y]$ yarıaçık aralığının büyüklüğü

$$dg(x, y] := g(y) - g(x)$$

olarak ilan edilir (Burada yarıaçık aralıklara geçtiğimize dikkat ediniz; çünkü birazdan göreceğimiz gibi g'nin süreksizlik noktalarındaki sıçramalar nedeniyle tek bir noktanın bile ağırlığı pozitif olabilir ve ardışık iki kapalı aralığın aradaki ortak noktada çakışıyor olması işleri karıştırabilir. Ardışık soldan yarıaçık aralıklar ise birbirinden ayrıktır, "ağırlıklar" birbirine karışmaz).

Bu durumda bu yeni "ağırlık" sistemine göre ağırlıklı toplamımız

$$\sum_{i=1}^{n} f(x_{i-1})(g(x_i) - g(x_{i-1}))$$

şeklindeki "Riemann-Stieltjes toplamı"dır. Dikkat edilirse, g'nin çok az arttığı aralıkların toplamda hemen hiç payı yokken g'nin hızlı artış gösterdiği aralıkların ağırlığı büyüktür. Yine gösterilebilir ki parçalanış incelirken bir limit gelir ve bu limit

$$\begin{aligned}\int_a^b f\,dg &:= \int_a^b f(x)\,dg(x)\\ &:= \lim_{n\to\infty} \sum_{i=1}^{n} f(x_{i-1})(g(x_i) - g(x_{i-1}))\end{aligned}$$

olarak yazılan Riemann-Stieltjes integralidir[2].

Hemen bazı temel özellikleri listeleyelim:

- Eğer $g(x) \equiv sabit$ ise her f için $\int_a^b f\,dg = 0$ olduğu kolayca görülebilir.
- Benzer şekilde $h = g + sabit$ ise $dh = dg$.
- İntegral f'ye göre doğrusaldır.
- $g(x) = x$ alırsak Riemann integralini elde ederiz. Yani Riemann-Stieltjes integrali, ilkinin bir genellemesidir.
- Eğer g türevli ve türevi sürekli bir fonksiyon ise, ince bir parçalanışta

  $$\begin{aligned}&\sum_{i=1}^{n} f(x_{i-1})(g(x_i) - g(x_{i-1}))\\ &\qquad\approx \sum_{i=1}^{n} f(x_{i-1})g'(x_{i-1})(x_i - x_{i-1})\end{aligned}$$

  olduğunu gözleyerek

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

  olduğunu tahmin edebiliriz ve bu doğru bir tahmindir. Yani bu durumda kalkülüsten alıştığımız $dg(x) = g'(x)dx$ değişimini yapabiliriz ve Riemann integraline dönmüş oluruz.

Bu yeni integralin getirdiği sürprizlere bir örnek verelim:

$$g(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

olsun. Sürekli bir $f : \mathbb{R} \to \mathbb{R}$ için $\int_{-1}^{1} f\,dg$ ne olur?

[2]Burada noktasal kütleler nedeniyle $[a, b]$ aralığının uç noktalarının hesaba dahil edilip edilmeyeceği sorusu dikkatli okuyucunun gözünden kaçmamıştır. Anlatımı dallandırıp budaklandırmamak için eksik kalan bu teknik ayrıntıyı görmezden geleceğiz.

Biz $[-1,1]$ aralığının herhangi bir $(x_i)_{i=0}^n$ parçalanışını aldığımızda sadece tek bir i_0 indisi için $0 \in (x_{i_0-1}, x_{i_0}]$ olacaktır ve

$$g(x_i) - g(x_{i-1}) = \begin{cases} 1, & i = i_0 \\ 0, & i \neq i_0 \end{cases}$$

olduğu hemen görülebilir. Bu durumda parçalanışın Riemann-Stieltjes toplamı

$$\sum_{i=1}^{n} f(x_{i-1})(g(x_i) - g(x_{i-1})) = f(x_{i_0-1})$$

çıkacaktır. Şimdi parçalanış incelirken $x_{i_0-1} \to 0$ olduğu için limite geçerek (ve f'nin sürekliliğini hatırlayarak)

$$\int_{-1}^{1} f\, dg = f(0)$$

elde ederiz.

İşte burada gördüğümüz olay, g'nin orijinde yarattığı sıçramanın o noktada noktasal kütle (point mass) oluşturmasıdır. Üstteki g o sıçrama noktası dışında sabit olduğu için bu noktasal kütle dışında tamamen "ölü" bir ölçüdür, yani 0'ı içermeyen tüm kümelere 0 ağırlık verir. Bu, aslında Ölçü Kuramı'nda δ_0 ile gösterdiğimiz, 0 noktasına 1 birim ağırlık biçen ve geri kalan her şeyin ağırlığını sıfır gören Dirac delta ölçüsünden başkası değildir.

Genel olarak g fonksiyonu p noktasında c birimlik sıçrama yapıyorsa, o noktada c birimlik noktasal kütle oluşur ($c\delta_p$ ölçüsü). Bir sonraki örnekle, bu davranışın okur zihninde netleşeceğini umuyoruz:

Örnek 5.1. Bir $g : \mathbb{R} \to \mathbb{R}$ fonksiyonunu

$$g(x) = \begin{cases} -1, & x < -2 \\ x+3, & -2 \leq x < 0 \\ x^2+7, & 0 \leq x < 1 \\ 9 & x \geq 1 \end{cases}$$

olarak tanımlayalım. Bu azalmayan fonksiyon -2, 0 ve 1 noktalarında sırayla 2, 4 ve 1 birimlik sıçramalar yapmaktadır ve aralarda türevlidir. Ayrıca sabitlikten ötürü $x < -2$ ve $x > 1$ bölgelerinde $dg = 0$. Dolayısıyla, sürekli bir $f : \mathbb{R} \to \mathbb{R}$ için

$$\begin{aligned} \int_{-3}^{2} f\, dg &= 2f(-2) + \int_{-2}^{0} f(x) d(x+3) \ + 4f(0) \\ &\quad + \int_{0}^{1} f(x) d(x^2+7) \ + 1 \cdot f(1) \\ &= 2f(-2) + \int_{-2}^{0} f(x) dx \ + 4f(0) \\ &\quad + \int_{0}^{1} 2x f(x) dx \ + f(1) \end{aligned}$$

olacaktır.

Azalmayan ama sürekli olup dolayısıyla hiçbir sıçraması olmayan Şeytan Merdiveni fonksiyonunu hatırlayın (sayfa 10). Eğer g olarak bu fonksiyonu alırsak Cantor kümesinin silinen açık aralıklarında $dg = 0$ olduğu görülür ve dg ölçüsünün desteği, yani "yaşadığı yer", tamı tamına Cantor kümesidir. Böylece burada hiçbir noktasal kütle içermemesine rağmen tüm kütlesi "sıfır uzunluklu" bir kümeye yığılmış bir ölçüyle karşı karşıyayız.

6 von Neumann'ın spektral temsil teoremi

Bu büyük sonucu alıntılamadan önce, ardındaki sezgiyi bir ölçüde yansıtacağını umduğumuz, sonlu boyuttan bir örnekle başlayalım:

Şimdi $\dim H < \infty$ ve $T : H \to H$ özeşlenik olsun. Elimizde T-özvektörlerinden oluşan bir ortonormal $(e_1, \ldots, e_n)$ tabanı olsun. Yani her $x \in H$ için

$$x = \sum_{i=1}^{n} \alpha_i e_i$$

yazabiliyoruz; burada elbette $\alpha_i = \langle x, e_i \rangle$ olmak zorundadır.

Bizim e_i özvektörlerimizin bazıları aynı özdeğere ait olabilir. Diyelim ki özdeğerlerimiz $\lambda_1 < \lambda_2 < \cdots < \lambda_k$, $k \leq n$ olsun. Üstteki toplamda aynı özdeğere ait terimleri gruplayarak

$$x = \sum_{i=1}^{k} v_i, \qquad Tv_i = \lambda_i v_i \tag{*}$$

şeklinde biricik bir temsil elde edebiliriz. Bu aslında H uzayının, birbirine dik T-özaltuzaylarına parçalanışıdır:

$$H = \bigoplus_{i=1}^{k} E_{\lambda_i} \qquad E_{\lambda_i} = \lambda_i\text{-özaltuzay.}$$

Şimdi bir dizi izdüşümü şöyle tanımlayalım: $i = 1, \ldots, k$ için

$$P_{\lambda_i} := \operatorname{span}\{E_{\lambda_j} : j < i\} \text{ üzerine izdüşüm}$$

ve $P_{\lambda_{k+1}} = I$ olsun. Yani (*) temsiline göre

$$\begin{aligned} P_{\lambda_1} x &= 0, \\ P_{\lambda_2} x &= v_1, \\ P_{\lambda_3} x &= v_1 + v_2, \\ &\vdots \\ P_{\lambda_k} x &= v_1 + v_2 + \cdots + v_{k-1}, \\ P_{\lambda_{k+1}} x &= v_1 + v_2 + \cdots + v_k = x \end{aligned}$$

olur. Demek ki

$$\begin{aligned} Tx = \sum_{i=1}^{k} Tv_i = \sum_{i=1}^{k} \lambda_i v_i &= \sum_{i=2}^{k+1} \lambda_{i-1} v_{i-1} \\ &= \sum_{i=2}^{k+1} \lambda_{i-1} (P_{\lambda_i} - P_{\lambda_{i-1}}) x \end{aligned}$$

yazabiliriz. Bu, bir Riemann-Stieltjes toplamına benzemektedir. Aslında her $\lambda \in \mathbb{R}$ için

$$P_\lambda := \text{span}\{E_{\lambda_i} : \lambda_i < \lambda\} \text{ üzerine izdüşüm}$$

tanımlarsak, λ'ya göre "artan", "soldan sürekli" ve özdeğerlerde ilgili özaltuzaya izdüşüm şeklinde sıçramaları olan bir $(P_\lambda)_{\lambda\in\mathbb{R}}$ dik izdüşüm ailemiz olur ve Riemann-Stieltjes integrali anlamında

$$Tx = \int_{\mathbb{R}} \lambda dP_\lambda x$$

yazabiliriz. Elbette burada sayı değerli toplamlardan vektör değerli toplamlara yapılacak bir uyarlama söz konusudur ancak okuyucunun aradaki geçişin yapılabileceğine ikna olduğunu umuyoruz.

Üstte T'yi n kere uygulayarak her $n \in \mathbb{N}$ için

$$T^n x = \int_{\mathbb{R}} \lambda^n dP_\lambda x,$$

dolayısıyla her polinom fonksiyon $f(\lambda)$ için

$$f(T)x = \int_{\mathbb{R}} f(\lambda) dP_\lambda x$$

temsili elde edilebilir. İşte von Neumann'ın teoremi, bu temsili sürekli f'lere ve hatta bunun biraz daha ötesine genellemektedir. Teoremin hemen öncesinde bir tanım verelim:

Tanım 6.1. Bir $(P_\lambda)_{\lambda\in\mathbb{R}}$ (dik) izdüşüm ailesine aşağıdaki koşulları sağlıyorsa bir *spektral aile* denir:

(i) $\lim_{\lambda\to-\infty} P_\lambda = 0$, $\lim_{\lambda\to+\infty} P_\lambda = I$ (limitler operatör normunda).

(ii) $\lambda \geq \mu$ ise $R(P_\lambda) \supseteq R(P_\mu)$ (yani izdüşülen altuzaylar büyür).

(iii) Soldan süreklilik: $\lim_{\mu\nearrow\lambda} P_\mu = P_\lambda$.

Teorem 6.2 (von Neumann'ın Spektral Temsil Teoremi)**.** *$T : D(T) \subseteq H \to H$ özeşlenik bir YATDD olsun. O halde T'ye karşılık bir $(P_\lambda)_{\lambda\in\mathbb{R}}$ spektral ailesi aşağıdaki koşulları sağlayacak şekilde bulunabilir:*

(i) T'nin tanım uzayı

$$D(T) = \left\{ x \in H : \int_{-\infty}^{\infty} \lambda^2 \langle dP_\lambda x, x\rangle < \infty \right\}$$

kümesi olur ve her $x \in D(T)$ için

$$Tx = \int_{-\infty}^{\infty} \lambda dP_\lambda x$$

sağlanır. Bu integraller Riemann-Stieltjes integrali olarak alınacaktır ve ikincisi H'nin normunda yakınsar. Ayrıca, dP_λ ölçüsünün desteği $\sigma(T)$ kümesidir ve $\lambda \mapsto dP_\lambda$ dönüşümünün süreksizlik noktaları tam olarak $\sigma_p(T)$ kümesidir (özdeğerler).

(ii) Eğer $f : \mathbb{R} \to \mathbb{C}$ sürekli ise bir $f(T)$ operatörü şu şekilde tanımlanabilir:

$$D(f(T)) = \left\{ x \in H : \int_{-\infty}^{\infty} |f(\lambda)|^2 \langle dP_\lambda x, x\rangle < \infty \right\}$$

olmak üzere $f(T) : D(f(T)) \subseteq H \to H$ dönüşümü her $x \in D(f(T))$ için H'nin normunda yakınsayan

$$f(T)x = \int_{-\infty}^{\infty} f(\lambda) dP_\lambda x$$

Riemann-Stieltjes integrali ile tanımlanır. Bu durumda

$$(f(T))^* = \overline{f}(T)$$

olacaktır; özellikle f gerçel değerliyse $f(T)$ özeşleniktir. Ayrıca

$$f(T) \text{ süreklidir} \iff f \text{ fonksiyonu } \sigma(T)\text{'de sınırlı ise.}$$

(iii) Eğer $f : \mathbb{R} \to \mathbb{C}$ Lebesgue ölçülebilir ve sınırlı ise $f(T)$ yukarıdaki gibi tanımlanabilir; ancak integral bir Lebesgue-Stieltjes integrali olarak alınır ve $f(T)x$ integrali H'de zayıf yakınsar, yani her $x \in D(f(T))$ ve $y \in H$ için

$$\langle f(T)x, y\rangle = \int_{-\infty}^{\infty} f(\lambda) \langle dP_\lambda x, y\rangle .$$

Biz elbette kanıtı burada vermeyeceğiz ama sürekli durumdaki kanıt stratejisine dair bir özet sunacağız. Buna başlamadan önce son kısımla ilgili bir not düşelim:

Bir X norm uzayından skaler cisme giden tüm sürekli doğrusal fonksiyonellerin uzayına X'in *dual uzayı* denir ve bu uzay X' ile gösterilir[3]. Bir X norm uzayındaki *zayıf topoloji*, her $f \in X'$ fonksiyonelini sürekli yapan en zayıf topoloji olarak tanımlanır. Eğer X'te bir $(x_n)_n$ dizisi bu topolojide bir x elemanına yakınsıyorsa buna zayıf yakınsama denir ve $x_n \xrightarrow{w} x$ yazılır. Şu önerme standart bir alıştırma sorusudur:

$$x_n \xrightarrow{w} x \quad \iff \quad \forall f \in X' \ f(x_n) \to f(x).$$

Eğer H bir Hilbert uzayı ise, Riesz Temsil Teoremi'nden biliyoruz ki H' aslında $y \leftrightarrow \langle \cdot, y\rangle$ eşleşmesi ile H'ye eşlenebilir. Bu durumda Hilbert uzayında zayıf yakınsama için

$$x_n \xrightarrow{w} x \quad \iff \quad \forall y \in H \ \langle x_n, y\rangle \to \langle x, y\rangle$$

yazabiliriz. Teoremin son kısmında bahsedilen yakınsama tam da budur. $f(T)x$'i tanımlamakta kullanılan Lebesgue-Stieltjes toplamları, bir limite zayıf yakınsamaktadır.

Şimdi kanıtla ilgili ipuçlarına dönelim. Önce bir teoremtanım:

Teorem 6.3. *Bir özeşlenik $T : D(T) \subseteq H \to H$ için şu önermeler denktir:*

(a) Her $x \in D(T)$ için $\langle Tx, x\rangle \geq 0$ (dikkat edilirse özeşleniklikten ötürü $\langle Tx, x\rangle \in \mathbb{R}$ olmak zorundadır).

[3] X^* gösterimi de yaygındır.

(b) $\sigma(T) \subseteq [0, \infty)$.

Bu koşulları sağlayan bir T dönüşümüne pozitif dönüşüm denir ve bu $T \geq 0$ ile gösterilir.

$$S \geq T :\iff S - T \geq 0$$

ile tanımlanan bağıntı, tüm H'de tanımlı (sürekli) özeşlenik dönüşümlerde bir kısmi sıralama bağıntısıdır. Ayrıca böyle T dönüşümleri için spektrum kümesinin minimal ve maksimal elemanları sırayla

$$\inf_{\|x\|=1} \langle Tx, x\rangle \quad ve \quad \sup_{\|x\|=1} \langle Tx, x\rangle$$

ile verilir.

Pozitiflik elde olunca, gerçel sayılarda yaptığımız bazı tanımların benzerleri özeşlenik dönüşümlere uyarlanabilir:

Teorem 6.4. *$T : H \to H$ (sürekli) özeşlenik olsun. Bu durumda T^2 pozitiftir ve $S^2 = T^2$ şartını sağlayan biricik bir pozitif $S : H \to H$ özeşlenik dönüşümü vardır ve bu $|T|$ ile gösterilir. Ayrıca,*

$$T^+ := \frac{|T| + T}{2}, \quad T^- := \frac{|T| - T}{2}$$

ile tanımlanan T^+ ve T^- dönüşümleri de (özeşlenik ve) pozitiftir ve T ile değişirler.

Bu son teoremin ilk kısmı aslında pozitif dönüşümlerin biricik pozitif karekökü olduğunu söylemektedir. Bunun kanıtı $f(z) = \sqrt{z}$ fonksiyonunun kesik düzlemdeki bir analitik koluyla yapılacağı gibi[4] tıpkı gerçel sayılarda bir sayıdan başlayıp o sayının kareköküne yakınsayan rekürsif tanımlı diziler kurulduğu gibi, bunu operatörlere uyarlayıp istenen kareköke yakınsayan bir dizi kurulabilir. Gerekli hesaplar zor olmasa da zaman alan uzun geçişler içerdiğinden kanıtı buraya almıyoruz.

Son olarak, (dik) izdüşümler hakkında söyleyeceklerimiz var. Bu kısımda geleneği bozmayıp izdüşümleri P ile göstereceğiz, bunun önceki bölümlerdeki momentum operatörüyle ilgisi yoktur:

Teorem 6.5. *$P : H \to H$ kapalı bir altuzaya dik izdüşümse P özeşlenik ve pozitiftir, ayrıca $P^2 = P$'dir. Eğer P' bir başka izdüşümse şunlar eşdeğerdir:*

(i) $P \geq P'$,

(ii) $R(P) \supseteq R(P')$,

(iii) $PP' = P'P = P'$.

Yani büyük altuzaya izdüşüm (operatör kısmi sıralamasında) daha büyüktür. Ayrıca üstteki durumda $P - P'$ de bir izdüşümdür ve görüntü uzayı, $R(P')$'nün $R(P)$ içindeki dik tümleyenidir[5].

[4] $0 \in \sigma(T)$ durumunda ufak bir cambazlık gerekmektedir!

[5] $Y \subseteq H$ bir kapalı altuzay ise $H = Y \oplus Y^\perp$ şeklinde parçalanır, burada $Y^\perp = \{x \in H : \forall y \in Y\ \langle x, y\rangle = 0\}$. Bu $Y^\perp$ uzayına Y'nin dik tümleyeni (orthogonal complement) denir.

Bu son cümleyi örneklemek için ℓ^2'de P ve P' izdüşümlerini

$$Px = (x_1, x_2, x_3, 0, 0, \ldots), \qquad P'x = (x_1, x_2, 0, 0, \ldots)$$

olarak alabiliriz. Görüldüğü üzere $P \geq P'$ ve $P - P'$ aslında span$\{e_3\}$ (üçüncü bileşen) altuzayına izdüşümdür ki bu tam da span$\{e_1, e_2, e_3\}$ içinde span$\{e_1, e_2\}$'nin dik tümleyenidir.

Şimdi, sürekli bir T için spektral ailemizi şöyle tanımlıyoruz:

$$P_\lambda := Ker((T - \lambda)^+) \text{ üzerine izdüşüm.}$$

Dikkat edilirse λ çok küçükse (spektrumun solundaysa) $T-\lambda$ pozitif ve birebirdir, yani $(T-\lambda)^+ = T-\lambda$ ve dolayısıyla çekirdek $\{0\}$ olduğundan izdüşüm de $P_\lambda = 0$ olur. Eğer λ çok büyükse bu sefer $(T - \lambda)^+ = 0$ ve çekirdek H olduğundan $P_\lambda = I$ olur. Genel olarak, λ büyüdükçe $T - \lambda$ negatifleşir, dolayısıyla pozitif kısmı küçülür ve pozitif kısmın çekirdeği büyür; yani P_λ'lar büyür (aslında bu şekilde kurduğumuz aile soldan değil sağdan sürekli olacaktır ama klasik numarayı yaparak P_λ yerine $P'_\lambda := \lim_{\mu \nearrow \lambda} P_\lambda$ alırsak bu sorun çözülür).

Temsil teoreminin kanıtında en kritik adımlarından biri şu gözlem olacaktır: Bir $\lambda \in \mathbb{R}$ ve küçük bir $d\lambda > 0$ için $dP_\lambda := P_{\lambda+d\lambda} - P_\lambda$ yazacak olursak

$$\lambda dP_\lambda \leq TdP_\lambda = (dP_\lambda)T \leq (\lambda + d\lambda)dP_\lambda$$

eşitsizliği doğrudur. Bu ne demektir? Şöyle diyebiliriz: dP_λ'nın izdüştüğü kısımda T'nin etkisi hemen hemen λ ile çarpım gibidir. Yani y vektörü dP_λ'nın izdüşüm uzayındaysa o zaman $y = dP_\lambda y$ olur ve $Ty = TdP_\lambda y \approx \lambda dP_\lambda y = \lambda y$ gibidir. Elbette bu mutlak anlamda değil ama $\langle Ty, y\rangle$ terimlerinin sıralanışı anlamında doğrudur, ama işin özü itibarıyla dP_λ "neredeyse" λ-özaltuzayına bir izdüşümdür. Burada λ bir özdeğer olmasa da izdüşülen kısımdaki vektörler yukarıdaki anlamda λ-özvektörmüş gibi davranırlar. Eğer λ gerçekten bir özdeğerse dP_λ'nın izdüştüğü uzay λ-özaltuzayını içerecektir. Daha önce momentum operatörünün fizikçi bakışıyla irdelemesinde kullandığımız, (6) ile verilen $\Delta\Psi$ vektörü aslında böyle bir izdüşüm bileşenini temsil etmektedir.

Üstelik $[\lambda_1, \lambda_2] \cap [\lambda'_1, \lambda'_2] = \emptyset$ ise $\Delta P = P_{\lambda_2} - P_{\lambda_1}$ ve $\Delta' P = P_{\lambda'_2} - P_{\lambda'_1}$ izdüşümleri birbirine dik altuzaylara izdüşeceklerdir. Biz esasen $f(T)$ dönüşümünü kurarken, x vektörünün $dP_\lambda x$ bileşenine $f(\lambda)$ ile çarpım gibi etkiyen bir dönüşüm inşa etmek istiyoruz. Eğer bu olacaksa $\mathbb{R}$'deki parçalanış inceldikçe

$$\sum_i f(\lambda_i) dP_{\lambda_i} x \tag{*}$$

toplamını yakınsak çıkartabilmemiz gerekir (ve bunun limit değeri $f(T)x$ olarak tanımlanacaktır). Üstteki ifadenin formal olarak normunun karesine bakalım:

$$\left\| \sum_i f(\lambda_i) dP_{\lambda_i} x \right\|^2 = \left\langle \sum_i f(\lambda_i) dP_{\lambda_i} x, \sum_j f(\lambda_j) dP_{\lambda_j} x \right\rangle.$$

Ortogonallikten, sağda yalnızca $i = j$ olan terim çiftleri kalır ve

$$\begin{aligned}\left\|\sum_i f(\lambda_i) dP_{\lambda_i} x\right\|^2 &= \sum_i \langle f(\lambda_i) dP_{\lambda_i} x, f(\lambda_i) dP_{\lambda_i} x\rangle \\ &= \sum_i |f(\lambda_i)|^2 \left\langle (dP_{\lambda_i})^2 x, x\right\rangle \\ &= \sum_i |f(\lambda_i)|^2 \langle dP_{\lambda_i} x, x\rangle\end{aligned}$$

gelir (burada izdüşümün özeşlenik olması ve kendi karesine eşit olmasını kullandık). Arzumuz bu toplamın yakınsamasıdır. Bu son kısım tam da von Neumann'ın teoremindeki Riemann-Stieltjes integraline karşılık gelir. Bu integralin yakınsadığı x'ler için (*) ifadesi limitte bize bir $f(T)$ tanımlar.

Sürekli durumu açıkladık ama Kuantum Mekaniği'ndeki Q ve P sürekli değildirler. Bu durumda ne yapılabilir?

Daha önce söylemedik ama, üniter dönüşümlerin spektrumu birim çember yani $\mathbb{S}^1$ içinde kalmak zorundadır. Spektral Dönüşüm Teoremi'nden ilham alarak, $\mathbb{R}$'yi $\mathbb{S}^1$'e taşıyan bir dönüşümün bir özeşlenik operatörün spektrumunu $\mathbb{S}^1$'e taşıyarak bize bir üniter operatör bırakması umulabilir. Gerçekten de

$$U := (T - i)(T + i)^{-1}$$

Möbius dönüşümü ile tanımlanan U operatörü (buna Cayley dönüşümü denir), T'nin YATDD (ve özeşlenik) olduğu durumda bile sürekli ve üniterdir, üstelik $1 \notin \sigma(U)$. Sürekli T'ler için elde edilen spektral temsil aracılığıyla üniter operatörlerin bir integral temsili bulunabilir ve

$$T = i(1 + U)(1 - U)^{-1}$$

ters Möbius dönüşümü ile süreksiz T durumunda bile bir integral temsili yazılabilir.

7 Ölçümlerin istatistiği, Heisenberg Belirsizlik İlkesi

Bu bölümle birlikte fiziğe geri dönüyoruz. Amacımız ilk olarak dalga fonksiyonunun olasılık yorumundan bahsetmek olacak, ardından Heisenberg'in Belirsizlik İlkesi'ne değineceğiz.

7.1 Olasılık ile ilgili önbilgiler

Bu kısımda işi Ölçü Kuramı'na girmeden halledebilmek adına kendimizi basit durumlarla sınırlayıp biraz yüzeysel kalacağız. Atladığımız bazı matematiksel ayrıntılar olacak; ama günün sonunda bunlar Kuantum Mekaniği örneklerimiz açısından gözardı edilebilirler.

Elimizde p_i olasılıkla λ_i sonucunu üreten $(i = 1, 2, \ldots)$ bir X "mekanizması" olsun. Olasılık kuramında bu, belli bir olasılık ölçüsü uzayından λ_i'leri içeren bir kümeye giden ve adına "rastgele değişken" (random variable) denen bir fonksiyona karşılık gelir. Üstte söylediğimiz özelliği

$$P[X = \lambda_i] = p_i$$

ile göstereceğiz (Burada bir kez daha P harfini ödünç alıp yeni bir anlamda kullanıyoruz; bunun özeşlenik türev dönüşümüyle bir ilgisi yoktur!). Elbette tüm olasılıkların toplamı 1 olmalıdır, yani $\sum_i p_i = 1$.

Bu rastgele değişkenin *beklenen değeri* ya da *beklentisi* (expectation), ürettiği sonuçların olasılıklarına göre ağırlıklı ortalamasıdır:

$$\mathbb{E}[X] = \sum_i \lambda_i P[X = \lambda_i] = \sum_i \lambda_i p_i.$$

Büyük Sayılar Kanunu (Law of Large Numbers) adıyla bilinen bir teorem, X deneyi bağımsız şekilde tekrarlandığı zaman gelen sonuçların ortalamasının beklentiden sapma olasılığının sıfıra gittiğini söyler. Örneğin adil bir para atışında yazı ile 1 lira kazanıyor, tura ile 3 lira kaybediyorsak, beklenti atış başına 1 lira kaybetmemizdir ve diyelim bir milyon kere atış yaparsak çok büyük olasılıkla bir milyon liraya yakın bir miktarda para kaybetmiş olacağız demektir.

Üstteki toplam bir sonsuz toplamsa yakınsak olmak zorunda değildir. Biz, toplamın yakınsadığı (yani "integrallenebilir") bir rastgele değişkenle çalıştığımızı varsayalım.

Beklenti X'in doğrusal bir fonksiyonudur:

$$\mathbb{E}[\alpha X + \beta Y] = \alpha\mathbb{E}[X] + \beta\mathbb{E}[Y].$$

Şimdi yine iki adil para atışı örneği düşünelim. Birinde yazı 1 lira kazanç, tura 1 lira kayıp olsun. Diğerinde ise yazı 1 milyon lira kazanç, tura 1 milyon lira kayıp olsun. Her iki örnekte de beklenti atış başına 0 liradır. Ancak beklentiler aynı olsa da ikincisi birincisine göre çok daha uzak uçlar arasında salınan sonuçlar üretir. Bu farkı ortaya koymaya yarayan ölçülerden biri *varyanstır*. Varyans, beklentiden sapmanın karesinin beklentisidir, yani

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Yine bu beklentiyi veren toplamın yakınsadığını varsayacağız. Aslında "beklentiden sapmanın beklentisi" olan $\mathbb{E}[|X - \mathbb{E}[X]|]$ değerini de düşünebilirdik ama kare fonksiyonu mutlak değer fonksiyonuna göre analitik açıdan çok daha iyi davranışlı olduğu için üstteki varyans tanımı çok daha kullanışlıdır. Varyansın karekökü de *standart sapma* olarak bilinir:

$$\sigma(X) := \sqrt{Var(X)}.$$

Kolayca görüleceği üzere:

- $X \equiv c =$ *sabit* ise (yani 1 olasılıkla c değerini alıyorsa) $\mathbb{E}[X] = c$.
- X (hemen her yerde) sabittir ancak ve ancak $Var(X) = 0$ ise.
- $Y = X + c$ (c sabit) ise $\mathbb{E}[Y] = \mathbb{E}[X] + c$ ve $Var(Y) = Var(X)$.

Üstteki özellikleri ve beklentinin doğrusallığını kullanarak bir hesap yapalım:

$$\begin{aligned}Var(X) &= \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2\mathbb{E}[X]X + (\mathbb{E}[X])^2] \\ &= \mathbb{E}[X^2] + \mathbb{E}[2\mathbb{E}[X] \cdot X] + \mathbb{E}[(\mathbb{E}[X])^2] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + (\mathbb{E}[X])^2 \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2.\end{aligned}$$

Sonuçta

$$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2,$$

ve özel olarak $\mathbb{E}[X] = 0$ ise $Var(X) = \mathbb{E}[X^2]$.

Rastgele değişkenler üstteki gibi ayrık değerler almak yerine diyelim tüm $\mathbb{R}$'de değerler alan "sürekli" rastgele değişkenler olabilirler. Kimi durumlarda bu gibi değişkenlerin istatistiğini bir "olasılık yoğunluk" fonksiyonu üzerinden betimlemek mümkündür; yani öyle bir $p(\lambda)$ fonksiyonu vardır ki

$$P[X \in [\lambda, \lambda + d\lambda]] \approx p(\lambda)d\lambda,$$

ya da daha doğru deyişle

$$\lim_{d\lambda \searrow 0} \frac{P[X \in [\lambda, \lambda + d\lambda]]}{d\lambda} = p(\lambda)$$

olur. Bu durumda X'in bir I aralığından sonuç verme olasılığı

$$P[X \in I] = \int_I p(\lambda)d\lambda$$

ile verilir. Dolayısıyla

$$\mathbb{E}[X] = \int \lambda P[X \in [\lambda, \lambda + d\lambda]] = \int_{\mathbb{R}} \lambda p(\lambda)d\lambda \qquad (13)$$

ile verilir.

7.2 Heisenberg'in Belirsizlik İlkesi

Önce dalga fonksiyonunun olasılık yorumuna bakacağız. İşe sonlu boyuttan başlayalım: $\dim H < \infty$ olmak üzere $T : H \to H$ özeşlenik olsun, T-özvektörlerden oluşan ortonormal bir $(e_1, \ldots, e_n)$ tabanı alalım.

Sonlu boyutlu bir Hilbert uzayını sonlu tane değer sistemleri modellemede kullanabiliriz; örneğin belli bir yönde spin değeri sadece $+1$ veya -1 olabilen bir parçacığın temsili iki boyutlu bir Hilbert uzayında yapılabilir.

Şimdi elimizde bir dalga fonksiyonu olsun:

$$\psi = \alpha_1 e_1 + \cdots + \alpha_n e_n. \qquad (14)$$

Burada $Te_i = \lambda_i e_i$ dersek, e_i vektörünün temsil ettiği durum, T ölçümünün kesin olarak λ_i değerini verdiği durum olarak yorumlanacaktır (tartışmayı basitleştirmek için her vektörün ayrı bir özdeğere ait olduğunu varsayalım, öyle değilse yine gruplama yaparak aynı noktaya varabiliriz). Herhangi bir ψ durumu ise üstteki gibi çeşitli olasılıkların birarada barındığı bir durum olacaktır. Peki (14) durumundaki ölçümün istatistiği nasıl bulunabilir?

Elbette bu matematiksel olarak tek yanıtı olmayan, açık uçlu bir sorudur. Fiziksel olarak ise deneylerle uyuşan yanıt neyse onu bulmak zorundayız. Neyse ki bu senaryoda hem deneylerin doğruladığı, hem de matematiksel olarak çok doğal duran bir yanıt var:

Biz esasen her bir e_i'nin tüm ψ içindeki payını bulmak istiyoruz, yani dik kenarlardan her birinin hipotenüs içindeki payını. Akla

$$\frac{|\alpha_i|}{|\alpha_1| + |\alpha_2| + \cdots + |\alpha_n|}$$

oranını önermek gelebilse de, matematik sezgisine sahip bir okuyucu sanıyoruz doğru seçeneğin

$$\frac{|\alpha_i|^2}{|\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_n|^2}$$

olması gerektiği konusunda bize katılacaktır. Hilbert uzayı geometrisinin "özünü yakalayan", Pisagor-Parseval özdeşliklerine göz kırpan bu ifade deneylerin de doğruladığı yanıttır.

Demek ki ψ normalize edilmiş olsaydı, yani $\sum |\alpha_i|^2 = 1$ olsaydı o zaman ölçümün λ_i çıkma olasılığı $|\alpha_i|^2$ olacaktı. Peki böyle bir ψ için T ölçümünün beklenen değeri nedir?

$$\mathbb{E}_\psi[T] = \sum_{i=1}^n |\alpha_i|^2 \lambda_i = \sum_{i=1}^n \langle \alpha_i \lambda_i e_i, \alpha_i e_i \rangle = \langle T\psi, \psi \rangle .$$

Son eşitlikte e_i'lerin ortogonalliğinin kullanıldığına dikkat ediniz.

Bunu sonsuz boyutlu Hilbert uzaylarındaki dalga fonksiyonlarına genelleyebiliriz. Normalize edilmiş bir ψ dalga fonksiyonuna sahip bir sistemde A ölçümünün beklenen değeri

$$\mathbb{E}_\psi[A] = \langle A\psi, \psi \rangle \qquad (\|\psi\| = 1)$$

ile verilecektir. Eğer dalga fonksiyonumuz A'nın ölçtüğü a büyüklüğü üzerinden $\psi(a)$ şeklinde temsil edilmişse, $A\psi(a) = a\psi(a)$ olacağından

$$\mathbb{E}_\psi[A] = \langle A\psi, \psi \rangle = \int a|\psi(a)|^2 da \qquad (\|\psi\| = 1)$$

olacaktır. Bu da demek oluyor ki ((13) ile karşılaştırınız!) ölçümde alınan a değerlerinin olasılık dağılımının yoğunluk fonksiyonu $|\psi(a)|^2$ fonksiyonudur. Ölçümün bir I aralığından gelme olasılığı da

$$P_\psi[A \in I] = \int_I |\psi(a)|^2 da$$

olacaktır.

Artık Heisenberg'in ünlü teoremini vermeye hazırız. İleride tekrar kullanacağımız bir kavramı tanımlayıp teoremi yazacağız:

Tanım 7.1. H uzayındaki iki özeşlenik dönüşüm olan A ve B'nin kuantum Poisson braketi

$$[A, B] = \frac{i}{\hbar}(AB - BA)$$

olarak tanımlanır. Burada A ve B birer YATDD olabilirler; bu durumda AB ve BA (ve ikisinin farkı) anlamlı oldukları en geniş altuzayda tanımlı olacaklardır. Fiziksel ölçümlere karşılık gelen tüm dönüşümler (ve bileşkeleri) en azından $C_0^\infty(\mathbb{R}^n)$ üzerinde tanımlı oldukları için hep yoğun altuzayda tanımlı dönüşümler ifade edeceklerdir. A ve B özeşlenik iken $[A, B]$'nin de özeşlenik olduğuna dikkat ediniz.

Teorem 7.2 (Heisenberg'in Belirsizlik İlkesi)**.** *Bir kuantum sisteminin normalize edilmiş dalga fonksiyonu ψ olsun ve birer ölçüme karşılık gelen A ve B özeşlenik dönüşümlerimiz olsun. Ayrıca $\psi \in D(A) \cap D(B)$, $A\psi \in D(B)$ ve $B\psi \in D(A)$ olduğunu varsayalım[6]. Bu durumda bu ölçümlerin standart sapmaları*

$$\sigma_\psi(A)\sigma_\psi(B) \geq \frac{\hbar}{2}|\langle [A, B]\psi, \psi\rangle|$$

eşitsizliğini sağlar[7].

Kanıt. Öncelikle, gerekirse $A \mapsto A - \mathbb{E}_\psi[A]$ ve $B \mapsto B - \mathbb{E}_\psi[B]$ değişimlerini yaparak $\mathbb{E}_\psi[A] = \mathbb{E}_\psi[B] = 0$ varsayabiliriz. Bu değişimin standart sapma ve braketi aynı bıraktığına dikkat ediniz. Böylece $\sigma_\psi(A) = \sqrt{\mathbb{E}_\psi[A^2]}$ olacaktır (B için de benzer).

Kanıt bilindik bir numaraya dayanmaktadır: Her $t \in \mathbb{R}$ için

$$\langle (A + itB)\psi, (A + itB)\psi\rangle \geq 0$$

olduğuna dikkat ediniz. Açarsak

$$\begin{aligned}
&\langle A\psi, A\psi\rangle + \langle A\psi, itB\psi\rangle + \langle itB\psi, A\psi\rangle + \langle itB\psi, itB\psi\rangle \\
=&\left\langle A^2\psi, \psi\right\rangle + -it\left\langle BA\psi, \psi\right\rangle + it\left\langle AB\psi, \psi\right\rangle + t^2\left\langle B\psi, B\psi\right\rangle \\
=&\mathbb{E}_\psi[A^2] + t\left\langle i(AB - BA)\psi, \psi\right\rangle + t^2\left\langle B^2\psi, \psi\right\rangle \\
=&\mathbb{E}_\psi[A^2] + \hbar t\left\langle [A, B]\psi, \psi\right\rangle + t^2\mathbb{E}_\psi[B^2] \\
=&Var(A) + \hbar t\left\langle [A, B]\psi, \psi\right\rangle + t^2 Var(B) \geq 0 \quad \forall t \in \mathbb{R}.
\end{aligned}$$

Son satırdaki t değişkenli ikinci dereceden polinomun tüm katsayılarının gerçel olduğunu gözleyiniz ($[A, B]$ özeşleniktir!). Diskriminantın pozitif olmaması gerekir, bu eşitsizlik de tam olarak teoremde öne sürülendir. ☐

Demek ki A ve B birbiriyle değişmiyorlarsa (yani $AB \neq BA$ ise) ψ durumunda bu ölçümler aynı anda kesin olarak yapılamayabilir; üstelik biri ne kadar kesin ise (küçük bir değer aralığına sıkışmış ise) diğeri o kadar belirsiz olacaktır. Kuantum Mekaniği'nde daima aynı anda yapılabilen ölçümler birbiriyle değişen operatörlerle verilenlerdir.

Bunun şöyle bir destekleyici argümanı da vardır: Aynı anda yapılacak ölçümün operatörü ne olacaktır? Hem AB hem BA eşit güçte adaylardır. O halde bu ikisinin hangisi olduğu fark etmemelidir, yani $AB = BA$ olmalıdır. Öte yandan ölçümler özeşlenik operatörlerle verilmek zorundadır. Ama zaten A ve B özeşlenik iken AB özeşleniktir ancak ve ancak $AB = BA$ ise! Tüm taşlar yerine oturmuş oluyor.

Peki bizim konum ve momentum operatörlerimiz için durum nedir? Bir $\psi(q)$ dalga fonksiyonu temsili üzerinden bakalım:

$$\begin{aligned}
(QP - PQ)\psi(q) &= q \cdot (-i\hbar\psi'(q)) - (-i\hbar(q\psi(q))') \\
&= -i\hbar q\psi'(q) + i\hbar\psi(q) + i\hbar q\psi'(q) = i\hbar\psi(q).
\end{aligned}$$

Demek ki

$$[Q, P] = -I.$$

Bu durumda herhangi bir normalize ψ için

$$\sigma_\psi(Q)\sigma_\psi(P) \geq \frac{\hbar}{2}|\left\langle -\psi, \psi\right\rangle| = \frac{\hbar}{2}.$$

Yani hiçbir durumda konum ve momentumu birlikte kesin olarak ölçmek mümkün değildir (çok boyutlu modelde aynı eksen boyunca olan konum ve momentum bileşenleri aynı anda ölçülemez, $i \neq j$ için $[Q_i, P_j] = 0$ olduğundan farklı bileşenlerin ölçümü aynı anda yapılabilir).

Aslında Q ve P ile ilgili bu durum bir başka yoldan da görülebilir: Konum ve momentum temsillerinin birbirine Fourier ve ters Fourier dönüşümleriyle dönüştüğünü hatırlayın. Konumun çok kesin biçimde ölçüldüğü bir durumda $\psi(q)$ dalga fonksiyonunun desteği (ya da kütlesinin çoğu) çok küçük bir bölgeye sıkışmış olacaktır. Ama Fourier Analizi'nden biliyoruz ki destek ne kadar lokalize olmuşsa (sıkışmışsa) transform o kadar dağınık olur; ters dönüşümde de aynı kural geçerlidir. Bunun en uç örneği olarak şunu verebiliriz: Kesin olarak λ ölçümünü belirten δ_λ ölçüsünün (artık L^2'den çıkıp dağılımlar uzayına geçtik) Fourier dönüşümü bir sabit çarpan farkıyla $\varphi(p) = e^{-i\lambda p}$'dir. Ama $|\varphi(p)|^2 \equiv 1$'dir, yani ölçülebilecek momentum değerleri tüm $\mathbb{R}$'ye düzgün dağılmıştır, yani momentum istatistiksel olarak tümüyle belirsizdir! Bu noktada, δ_λ aslında bir L^2 fonksiyonu ile verilmediğinden fiziksel modelimiz açısından geçerli bir dalga fonksiyonu değildir.

7.3 Çok bileşenli sistemlerin temsiline ilişkin bir ek

Bu kısımda kısacık bir özet halinde birden çok bileşeni olan sistemlerin temsiline değineceğiz (örneğin iki elektron, ya da 3 "kubit"lik bir kuantum bilgisayar belleği).

Eğer ilk bileşenin durumu H_1, ikincisininki de H_2 uzayında temsil ediliyorsa, bileşik sistemin durumu $H_1 \otimes H_2$ tensör çarpımında temsil edilecektir ($k > 2$ bileşen için de yine k'lı bir çarpım kullanılır). İki vektör uzayı X ve Y'nin tensör çarpımını (teknik tanıma girmeden, sözlerle) şöyle tanımlayabiliriz: $X \otimes Y$ uzayı, $x \otimes y$ tipindeki elemanlarla gerilen ve elemanları

$$\alpha_1 x_1 \otimes y_1 + \cdots + \alpha_n x_n \otimes y_n, \quad x_i \in X, y_j \in Y, \alpha_i \in \mathbb{C}$$

şeklindeki sonlu formal toplamlardan oluşan bir vektör uzayı olacaktır. Bu formal toplamda sıralama önemli değildir

[6] Yukarıda açıkladığımız dalga fonksiyonunun olasılık yorumundan görüleceği üzere, çoğu uygulamada dalga fonksiyonları sonsuzda hızla sönen fonksiyonlar olduğu için P, Q ve bunlardan türetilmiş dönüşümler altında görüntüleri yine sonsuzda sönümlenen ve bu dönüşümlerin tanım kümesine düşen fonksiyonlar olmaktadır.

[7] Geleneği bozmayarak standart sapmayı σ ile gösterdik. Bunun spektrum ile karıştırılmayacağını umuyoruz.

(farklı sırada yazılmış toplam ilkiyle aynı kabul edilecektir) ve ayrıca şu denklikler tanımlanacaktır:

(i) $0x \times y = 0$ (boş toplam),

(ii) $\alpha(x \otimes y) = (\alpha x) \otimes y = x \otimes (\alpha y)$,

(iii) $(\alpha x_1 + \beta x_2) \otimes y = \alpha x_1 \otimes y + \beta x_2 \otimes y$,

(iv) $x \otimes (\alpha y_1 + \beta y_2) = \alpha x \otimes y_1 + \beta x \otimes y_2$.

Gösterilebilir ki $X \times Y$ üzerindeki her bilineer T dönüşümü için $X \otimes Y$'de tanımlı ve $T(x, y) = S(x \otimes y)$ şartını sağlayan biricik bir lineer S dönüşümü vardır; bu özellik tensör çarpımının karakteristik özelliğidir.

Eğer dalga fonksiyonları ψ_1 ve ψ_2 olan iki bileşen birbirinden "bağımsız" ise bunların birleşik dalga fonksiyonu $\psi_1 \otimes \psi_2$ olacaktır; daha doğrusu bunu bağımsız olmanın tanımı olarak alabiliriz. Böyle bir tanımın ardındaki nedeni, dalga fonksiyonunun olasılık yorumu üzerinden açıklamaya çalışalım:

Örneğimizde $H_1 = H_2 = \mathbb{C}^2$ olsun ve (e_1, e_2) de belli bir ölçümün özvektörlerinden oluşan bir ortonormal taban olsun. Dalga fonksiyonlarımız normalize edilmiş olarak

$$\psi_1 = ae_1 + be_2 \quad \text{ve} \quad \psi_2 = ce_1 + de_2$$

olsunlar, yani $a^2 + b^2 = c^2 + d^2 = 1$. Parçacıkların (isimleri x ve y olsun) bu ölçümde e_1 halinde bulunma olasılıkları[8] sırayla a^2 ve c^2 olacaktır, e_2 için de bunlar b^2 ve d^2'dir. Bağımsızlık durumunda birleşik sistemde bu istatistiklerin diğer tanecikten bağımsız şekilde korunmasını bekleriz. Şimdi eğer birleşik sistemde dalga fonksiyonu $\psi_1 \otimes \psi_2$ ise bunu

$$ace_1 \otimes e_2 + ade_1 \otimes e_2 + bce_2 \otimes e_1 + bde_2 \otimes e_2$$

olarak açabiliriz, elbette bu da normalize bir dalga fonksiyonudur:

$$(ac)^2 + (ad)^2 + (bc)^2 + (bd)^2 = (a^2 + b^2)(c^2 + d^2) = 1.$$

Şimdi bu sistemde örneğin y taneciği e_2 halindeyken x'in e_1 halinde olma olasılığı $e_1 \otimes e_2$'nin olasılığıdır, ki o da

$$(ad)^2 = a^2 d^2 = P(x = e_1) \cdot P(y = e_2)$$

olarak yazılabilir, yani tam da bağımsız değişkenlerin sağlaması gereken ilişki.

Klasik Mekanik'te arada mekanik bir etkileşim olmayan parçacık çiftleri bağımsızdır; ancak üstteki kuantum modelinde, herhangi bir mekanik etkileşim olmasa da bileşenlerin bağımsız olamadığı durumlar olabilir. Bunlar (iki bileşenli bir sistemde) tam da $\psi_1 \otimes \psi_2$ şeklinde çarpanlara ayrılamayan hallerdir (bu durumda parçacıklar *dolanıktır* denir). Buna bir örnek

$$\psi = \frac{1}{\sqrt{2}} e_1 \otimes e_1 + \frac{1}{\sqrt{2}} e_2 \otimes e_2$$

halidir. Bu sistem %50 olasılıkla $e_1 \otimes e_1$, %50 olasılıkla da $e_2 \otimes e_2$ halindedir; örneğin eşit olasılıkla belli bir yöndeki polarizasyonu ikisinin de $+1$ veya ikisinin de -1 ölçülen iki foton, ya da birlikte 0 veya birlikte 1 ölçülen iki kubit. Bu tür dolanık parçacık çiftleri deneylerde gerçekten oluşturulabilmektedir. Kuantum bilgisayarlarının ve kuantum haberleşme tekniklerinin temeli bunlara dayanır.

[8] Yani e_1'e karşılık gelen ölçümün alınma olasılığı.

8 Schrödinger denklemi

Şu ana kadar hep belli bir anda verilmiş dalga fonksiyonları üzerine konuştuk. Ama fizik elbette sistemin zamanla evrimini de incelemek zorundadır. Artık zaman parametresini de işin içine katacağız. Sonunda sistemin evrimini açıklayan Schrödinger Denklemi'ni elde edeceğiz.

8.1 Klasik Mekanik'in Hamilton formülasyonu

Bilindiği gibi Newton'un kanunlarına dayanan Klasik Mekanik'te yeni bir bakış açısı sunup problemlerin çözümünde yeni olanaklar sunan iki önemli formalizm vardır: Bunlardan ilki Varyasyonlar Kalkülüsü'ne (Calculus of Variations) dayanan Lagrange formalizmi, diğeri ise daha diferansiyel-geometrik olan Hamilton formalizmidir. Kuantum Mekaniği'nin temel denklemi olan Schrödinger Denklemi, klasik karşılığı bu ikincisi olan bir yaklaşıma dayanır. Bu nedenle önce Hamilton'un Klasik Mekanik formülasyonunu özetlemek istiyoruz.

Aslında bu yaklaşım en genel haliyle yüzeylerde diferansiyel formları ele alınarak sunulabilir; ancak bizim amacımız için aşağıdaki en temel model yeterli olacaktır:

Hamilton'ın formülasyonunda bir klasik sistemin durumu her bileşenin (tanecik veya tek bir bütün olan alınan bir kütle) konum ve momentum değerleriyle ifade edilir. Örneğin tek tanecikli bir boyutlu sistemde bir $(q, p) \in \mathbb{R}^2$ ikilisi, ya da üç boyutta bir tanecik için $(q_1, q_2, q_3, p_1, p_2, p_3) \in \mathbb{R}^6$ altılısı o taneciğin hareketine dair tüm bilgiyi içerir. Bu vektörler bizim ψ dalga fonksiyonları gibi sistem durumunu temsil etmektedir.

Fiziksel büyüklüklerin ölçümü ise bu vektörlerin bileşenlerinin fonksiyonlarıyla verilir. Yalınlık açısından bundan sonra örneklerimizi bir boyut üzerinden vereceğiz. Bu durumda her ölçüm işlemi $f(q, p)$ gibi bir fonksiyona karşılık gelir[9], örneğin kinetik enerji m kütle olmak üzere $f(q, p) = p^2/2m$ ile verilir. Bunlar da kuantum bakışındaki özeşlenik dönüşümlere karşılık gelen nesnelerdir.

Şimdi işin içine zamanı katalım. Başlangıçta (q_0, p_0) durumunda olan taneciğin durumu t kadar zaman sonra ne olur? Bu durumu $(q(t), p(t))$ ile gösterirsek, sistemin evrimi, adına faz uzayı denen qp düzleminde dolaşan parametrik bir $(q(t), p(t))$ eğrisi ile temsil edilebilir. Peki bu eğrinin sağladığı denklemler nedir?

Sistemin toplam enerjisini belirten ve adına Hamiltonyen denilen bir $H(t, q, p)$ fonksiyonu olacak. Biliyoruz ki

$$H = E_{toplam} = E_{kinetik} + E_{potansiyel}.$$

Burada H'nin t'ye bağlı olmasına izin verdik; çünkü sistemin potansiyel fonksiyonu zamanla değişebilir (bir deneycinin sisteme değişken bir elektrik alan uyguladığını düşünün). En genel haliyle

$$H(t, q, p) = \frac{p^2}{2m} + U(t, q) \tag{15}$$

[9] Aslında $f(t, q, p)$ daha doğru olurdu; bir sonraki kısmın sonundaki ek bilgiye de bakınız.

olacaktır, buradaki ikinci terim zamana ve uzaydaki konuma bağlı potansiyel fonksiyonudur.

Bir an için potansiyelin zamandan bağımsız olduğunu düşünelim. Bu durumda enerjinin korunum kanununa göre toplam enerji sabit kalmalıdır. Bu nedenle taneciğimiz qp düzleminde $H = sabit$ eğrilerini izlemelidir, bu da yörüngesi ∇H vektörüne olacak demektir.

İşte Hamilton'un denklemleri de tam buna karşılık gelir. Bu denklemler

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}, \qquad \frac{dp}{dt} = -\frac{\partial H}{\partial q} \tag{16}$$

olması gerektiğini söyler (yüksek boyutlu modellerde aynı indisli her bir q_i-p_i çifti için bu denklemler yazılır). Üstelik bu, zamana bağlı potansiyel durumunu da kapsar.

Hamilton denklemlerinin Newton yasalarına denk olduğunu görelim. Yukarıdaki (15) ve (16) ifadelerine göre

$$\frac{dq}{dt} = \frac{\partial}{\partial p}\left(\frac{p^2}{2m} + U(t,q)\right) = \frac{p}{m}$$

ki bu son ifade olması gerektiği gibi v, yani hızdır. Diğer denklem ise

$$\frac{dp}{dt} = -\frac{\partial}{\partial q}\left(\frac{p^2}{2m} + U(t,q)\right) = -\frac{\partial U}{\partial q}$$

olur ki bir potansiyel alanının uyguladığı kuvvet, o potansiyelin en hızlı düştüğü yönde olup potansiyel fonksiyonunun gradyanının eksilisiyle verilir. Böylece

$$\frac{dp}{dt} = F$$

şeklindeki Newton Kanunu'nu elde etmiş oluyoruz. Kütle sabitken bu $\frac{d(mv)}{dt} = m\frac{dv}{dt} = ma$ şeklini alır, yani ünlü $F = ma$ yasası!

Peki bu durumda yapılan bir f ölçümü zamanla nasıl değişir? Örneğin f hızı ölçüyorsa biz ivme ölçümünün q ve p cinsinden ifadesini aramış oluyoruz. Bunu da hesaplayalım:

$$\frac{df(q(t),p(t))}{dt} = \frac{\partial f}{\partial q}\frac{dq}{dt} + \frac{\partial f}{\partial p}\frac{dp}{dt} = \frac{\partial f}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial H}{\partial q} =: \{H, f\}. \tag{17}$$

Yukarıda tek boyutlu örneği görülen ve çok boyutta iki f ve g fonksiyonu için

$$\{g, f\} := \sum_i \left(\frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right)$$

şeklinde tanımlanan işleme *Poisson braketi* denir. Bazı özellikleri şöyledir:

- $\{f, g\} = -\{g, f\}$.
- Bilineerdir.
- Jacobi özdeşliğini sağlar:

  $$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0.$$

- Bir tür çarpım kuralını sağlar ($\{f, \cdot\}$ bir "derivation" işlemidir):

  $$\{f, gh\} = g\{f, h\} + \{f, g\}h.$$

- $f(\vec{q}, \vec{p}) = q_i$ ve $g(\vec{q}, \vec{p}) = p_i$ ile verilen konum ve momentum bileşen fonksiyonları

  $$\{q_i, q_j\} = \{p_i, p_j\} = 0, \qquad \{q_i, p_j\} = \delta_{ij}$$

  bağıntılarını sağlar.

8.2 Kuantum Mekaniği'nde durum

Bir kuantum sisteminin zamanla evrimi incelenirken iki yaklaşım vardır:

(i) Heisenberg resmi: Bu modelde dalga fonksiyonu $\psi(q)$ sabittir, buna karşın bir a büyüklüğünün ölçüm operatörü $A(t)$ zamana bağlıdır. Yani $t = 0$ referans anı ise, t kadar zaman sonraki ölçümü veren operatör $A(t)$'dir (ve $\psi(q)$ dalga fonksiyonu üzerine etkiyecektir).

(ii) Schrödinger resmi: Bu bakışta dalga fonksiyonu $\psi(q, t)$ zamana bağlıdır ve zamanla evrilmektedir, ölçüm operatörleri ise sabittir.

Önce bu yaklaşımların ufak bir felsefi tartışmasını yapalım: Schrödinger'in bakışının oldukça akla yatkın ve kabulünün kolay olduğuna okuyucunun da katılacağını düşünüyoruz. Sonuçta dalga fonksiyonu sistemin durumunu temsil etmektedir ve onun zamanla değiştiğini düşünmek doğaldır. Bu değişimin kesin bir kanuna dayandığını varsaymak da makul bir hipotezdir.

Öte yandan, bir ölçümün kendisi deterministik olmasa da ölçümün istatistiği dalga fonksiyonu ile belirlenir ve kesin olarak bellidir. İleride yapılacak bir ölçümün de sonucu önceden belli olmasa bile t zaman sonra yapılacak ölçümlerin istatistiği de şimdiden bilinebilir; çünkü bu istatistik deterministik bir şekilde evrilmektedir. Demek ki t zaman sonraki ölçüm hemen şimdi, şu anki dalga fonksiyonu üzerinden irdelenebilir. Üstteki anlatıda $A(t)$'nin rolü budur.

Aslında her iki bakış da Klasik Mekanik'teki Hamilton yaklaşımına dayanır. Schrödinger resmi, önceki kısımda özetlediğimiz ve adına Liouville resmi denilen $(q(t), p(t))$ "evrilen durum" bakışıdır. Buna karşılık

$$f_t(q, p) \longleftrightarrow f(q(t), p(t))$$

ile sembolize edebileceğimiz, durumu sabit kabul edip ölçüm fonksiyonlarının zamana bağlı bir $(f_t)_t$ dönüşüm ailesi olarak düşünen ve adına Hamilton resmi denen bir yaklaşım daha vardır. Bu ikisi birbirine denktir. İşte Heisenberg yaklaşımı bu klasik modelin karşılığıdır.

Şimdi biz, şu iki soruya yanıt arayacağız: Diyelim ki A hız ölçümü yapan dönüşüm olsun. Peki o zaman (i) $A(t)$ nedir?, (ii) hızın zamana göre türevi olan ivmeyi ölçen B dönüşümü nedir? İkincisinin Klasik Mekanik'te Poisson braketi (17) ile elde edilebileceğini biliyoruz.

Schrödinger resmine göre irdeleme yapalım. Dalga fonksiyonunun zamanla evrimini veren dönüşümün normalizasyonu korumasını isteriz. Bunu garanti etmek için de evrim dönüşümünün üniter olduğu hipotezini ortaya atacağız. Dolayısıyla bir $S(t), t \in \mathbb{R}$ üniter dönüşüm ailesi[10],

$$\psi(q,t) = S(t)\psi(q,0) := S(t)\psi(q)$$

şeklinde dalga fonksiyonunu evriltecektir. Üniterlikten ötürü $(S(t))^*$ için $S(t)^*$ yazarak)

$$S(t)S(t)^* = S(t)^*S(t) = 1 \tag{18}$$

olacaktır.

Şimdi de Heisenberg resmine dönersek, A ölçümü için bizi Heisenberg bakışındaki $A(t)$ dönüşümüne bağlayan diyagram aşağıdaki gibi olacaktır:

$$\begin{tikzcd}
\psi(q,t) \arrow[r, "A"] & A\psi(q,t) \arrow[d, "S(t)^*"] \\
\psi(q,0) \arrow[u, "S(t)"] \arrow[r, "A(t)"] & A(t)\psi(q,0)
\end{tikzcd}$$

Dolayısıyla

$$A(t) = S(t)^*AS(t).$$

Burada türev alarak ve çarpım kuralını uygulayarak

$$\frac{dA(t)}{dt} = \frac{dS(t)^*}{dt}AS(t) + S(t)^*A\frac{dS(t)}{dt} \tag{19}$$

elde ederiz. Bu noktada fiziksel tutarlılık açısından

$$\frac{dA(t)}{dt} = B(t)$$

olması gerektiğini öne sürüyoruz. Bu kabul edildikten sonra, üstteki diyagrama göre

$$B = S(t)B(t)S(t)^*$$

olacaktır ve bunu (19) eşitliğinde yerine koyarsak

$$\begin{aligned}
B &= S(t)\frac{dS(t)^*}{dt}AS(t)S(t)^* + S(t)S(t)^*A\frac{dS(t)}{dt}S(t)^* \\
&\overset{(18)}{=} S(t)\frac{dS(t)^*}{dt}A + A\frac{dS(t)}{dt}S(t)^*
\end{aligned}$$

gelir. Son olarak, yine (18)'de türev alarak

$$\frac{dS(t)}{dt}S(t)^* + S(t)\frac{dS(t)^*}{dt} = 0$$

elde eder ve bunu üstte kullanırsak

$$\tilde{H} = i\hbar\frac{dS(t)}{dt}S(t)^*$$

olmak üzere ($\tilde{H}$'nin özeşlenik olduğuna dikkat ediniz!)

$$B = \frac{i}{\hbar}(\tilde{H}A - A\tilde{H}) = [\tilde{H}, A]$$

elde ederiz. Bu son kısımdaki işlemin kuantum Poisson braketi olduğunu hatırlatıyoruz. Demek ki Klasik Mekanik'te olduğu gibi ölçülen büyüklüğün zamana göre değişim hızını veren ölçüm, yine bir braket işlemiyle elde ediliyor. Peki bu $\tilde{H}$ aslında nedir?

Zamandan bağımsız potansiyel enerji durumunda toplam enerji H'nin sabit kalması $[\tilde{H}, H] = 0$, yani $\tilde{H}H = H\tilde{H}$ olmasını gerektirir. Bu değişmeli olma durumu akla $\tilde{H} = f(H)$ formunda bir bağıntı aramayı getirir. Nitekim deneyler $\tilde{H} = H$ olduğunu göstermektedir. Sonuç olarak

$$B = [H, A].$$

Artık Schrödinger Denklemi'ni elde etmeye hazırız: Schrödinger resmine göre

$$\psi(q,t) = S(t)\psi(q,0)$$

idi. O halde türev alırsak

$$\begin{aligned}
\frac{\partial\psi(q,t)}{\partial t} &= \frac{dS(t)}{dt}\psi(q,0) = \frac{dS(t)}{dt}S(t)^{-1}\psi(q,t) \\
&= \frac{dS(t)}{dt}S(t)^*\psi(q,t) = \frac{1}{i\hbar}H\psi(q,t)
\end{aligned}$$

yani

$$H\psi(q,t) = i\hbar\frac{\partial\psi(q,t)}{\partial t} \qquad \text{(Schrödinger Denklemi)}$$

elde edilir. Operatör kalkülüsü ile buradan çözümün

$$\psi(q,t) = e^{-\frac{i}{\hbar}Ht}\psi(q,0)$$

olması gerektiği de görülebilir, yani böylece $S(t)$'yi de bulmuş oluyoruz:

$$S(t) = e^{-\frac{i}{\hbar}Ht}.$$

Tüm bunların üzerine, (3 boyutta) Klasik Mekanik'te olduğu gibi

$$[Q_i, Q_j] = [P_i, P_j] = 0, \qquad [Q_i, P_j] = \delta_{ij}$$

olduğunu da ekleyelim.

Örneklere geçmeden önce üstteki modelde ufak bir genelleme yapılabileceğini belirtelim: Bir A ölçümünün kendisi aslında t parametresine bağlı olabilir. Bununla $A(t)$ evrimini kastetmiyoruz; herhangi bir andaki ölçümün değeri zamana bağlı olabilir; örneğin zamana göre değişen bir potansiyel altında enerji ölçümü. Bu durumda (19) eşitliğinde ortada bir

$$S(t)^*\frac{\partial A}{\partial t}S(t)$$

terimi eklenecektir ve B ölçümü de

$$B = \frac{\partial A}{\partial t} + [H, A]$$

şeklini alacaktır. Bunun Klasik Mekanik karşılığı da $f = f(t, q(t), p(t))$ olmasına izin vermektir.

[10] Aslında bunlar tek parametreli bir grup oluştururlar.

8.3 Örnekler

Örnek 8.1 (Potansiyel altında serbest parçacığın enerjisi)**.** Kütlesi m olan bir parçacığımızın dalga fonksiyonu $\psi(q)$ ile verilsin. Bu parçacık bir $u(q)$ potansiyeli altında ise potansiyel ölçümünü yapan dönüşüm

$$U : \psi(q) \mapsto u(q)\psi(q)$$

şeklinde verilir ve dolayısıyla toplam enerji operatörü

$$H = \frac{P^2}{2m} + U$$

olacaktır. Ama konum temsilinde $P = -i\hbar\frac{\partial}{\partial q}$ olduğundan $P^2 = -\hbar^2\frac{\partial^2}{\partial q^2}$ gelir, buradan da (yüksek boyutta da geçerli olan)

$$H\psi = -\hbar^2\Delta\psi + u\psi$$

elde edilir.

Örnek 8.2 (Hız ölçümü operatörü)**.** Konumun Q tarafından ölçüldüğünü biliyoruz. O halde hız ölçüm operatörü $V = [H, Q]$ ile verilmelidir. Önceki örneğin şartlarında yine konum temsili üzerinden, tek boyutta gerekli hesabı yapalım:

$$\begin{aligned}
(V\psi)(q) &= ([H,Q]\psi)(q) \\
&= \frac{i}{\hbar}\left(\left(\frac{P^2}{2m} + U\right)Q\psi - Q\left(\frac{P^2}{2m} + U\right)\psi\right)(q) \\
&\overset{UQ=QU}{=} \frac{i}{2m\hbar}\left(P^2Q\psi - QP^2\psi\right)(q) \\
&= \frac{1}{2m\hbar}\left(-\hbar^2(q\psi(q))'' - (-\hbar^2 q\psi''(q))\right) \\
&= \frac{-i\hbar}{m}\psi'(q) = \left(\frac{P}{m}\psi\right)(q).
\end{aligned}$$

Öyleyse $V = \frac{P}{m}$, bu tam da Klasik Mekanik'teki ilişkidir!

Örnek 8.3 (Momentumun zamana göre türevi)**.** Yukarıdakine benzer şekilde momentumun zamana göre türevini veren dönüşüm $F = [H, P]$ ile hesaplanabilir. Bu durumda H içindeki $P^2/2m$ terimi P ile değiştiği için

$$F = \frac{i}{\hbar}(UP - PU)$$

olacaktır. Biraz hesapla bir boyutta $F\psi = -u'\psi$ ve yüksek boyutta

$$F_i\psi = -\frac{\partial u}{\partial q_i}\psi$$

olduğu görülür; bu da klasik kuramdaki $F = \frac{dp}{dt} = -\nabla u$ denkleminin karşılığıdır.

Örnek 8.4 (Kuantum harmonik salınıcı)**.** Bir denge durumunun etrafında kararlı bir şekilde salınmakta olan ve m kütleli bir tanecik düşünelim (bir elektron, atom vb.). Böyle bir tanecik sabit enerjisi nedeniyle bir H-özvektör durumunda olmalıdır ve dalga fonksiyonunun özdeğeri de parçacığın sabit $E > 0$ enerjisi olacaktır. Burada salınımın denge durumundaki ($q = 0$ alıyoruz) potansiyel 0 ve kütle/yay modelindeki gibi ω_0 frekanslı salınım halinde potansiyel enerji $\frac{1}{2}m\omega_0 Q^2$ alınacaktır. Dolayısıyla (tek boyutta) denklemimiz sabit bir E için

$$H\psi(q) = E\psi(q), \quad H = \frac{P^2}{2m} + \frac{1}{2}m\omega_0^2Q^2.$$

Bunlar yerine yazıldığında

$$\frac{-\hbar^2}{2m}\psi''(q) + \frac{1}{2}m\omega_0^2q^2\psi(q) = E\psi(q)$$

şeklini almaktadır. Şimdi

$$x = \sqrt{\frac{m\omega_0}{\hbar}}, \quad \lambda = \frac{E}{\hbar\omega_0}$$

değişken değişimlerini yaparsak denklemimiz

$$\psi''(x) - x^2\psi(x) = -2\lambda\psi(x)$$

haline gelir. Burada ortadaki terim bizi $\psi(x) = e^{-x^2/2}\varphi(x)$ değişimine yönlendirir ve bu yapıldığında yeni denklem

$$\varphi''(x) - 2x\varphi'(x) + (2\lambda - 1)\varphi(x) = 0$$

olur. Denklemin $\varphi(x) = \sum_{n=0}^{\infty} a_n x^n$ şeklinde bir kuvvet serisi çözümü arandığında

$$a_{n+2} = \frac{(2\lambda - 2n - 1)}{(n+1)(n+2)}a_n, n \geq 1, \quad a_0, a_1 \in \mathbb{R}$$

elde ederiz.

Burada şu kritik gözlemi yapıyoruz: Eğer her n için $2\lambda - 2n - 1 \neq 0$ ise aşikar çözüm $\varphi = 0$ dışında tüm çözümlerde tek veya çift indisli katsayılar için

$$\frac{a_{n+2}}{a_n} \sim \frac{2}{n}$$

olur, sağdaki ise tam da $e^{x^2} = \sum_{n=0}^{\infty}\frac{x^{2n}}{n!}$ serisinde görülen durumdur. Buna dayanarak, böyle bir serinin e^{x^2}'ye yakın hızda büyüyen bir fonksiyon vereceği ve $\psi(x) = e^{-x^2/2}\varphi(x) \notin L^2$ olacağı görülür. Halbuki fiziksel olarak anlamı olan çözümlerde $\psi(x) \in L^2$ olmalıdır. Bu da ancak seride sadece tek veya sadece çift terimler olup bir n için

$$2\lambda - 2n - 1 = 0$$

olursa mümkündür; bu durumda seri aslında adına Hermite polinomu denilen bir polinom olacaktır. Burada çözümün son halini yazmadan vurgulamak istediğimiz çarpıcı gerçek şudur: Bir $n \in \mathbb{N}$ için

$$\lambda = \frac{E}{\hbar\omega_0} = n + \frac{1}{2}$$

yani $E = \hbar\omega_0(n + \frac{1}{2})$ olmak zorundadır. Dolayısıyla klasik sistemin aksine kararlı salınım yapan bir taneciğin sahip olabileceği enerjiler kesiklidir!

Kaynaklar

[1] Kreyszig, Erwin. *Introductory functional analysis with applications*. Vol. 1. New York: Wiley, 1978.

[2] Schechter, Martin. *Principles of functional analysis*. No. 36. American Mathematical Society, 2001.

[3] Helmberg, Gilbert. *Introduction to spectral theory in Hilbert space*. Courier Dover Publications, 2008.

[4] Faddeev, Ljudvig D., ve Oleg Aleksandrovich Yakubovskiĭ. *Lectures on quantum mechanics for mathematics students*. Vol. 47. American Mathematical Society, 2009.

[5] Fok, V. A. *Principles of quantum mechanics*. MoIzN (1976).

[6] von Neumann, J. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1955.

[7] Takhtajan L.A. *Quantum Mechanics for Mathematicians*, American Mathematical Society, 2008.