L^2 空間とルベーグ積分

量子力学でたまに L^2 空間が出てくる理由を示します。数学側の話なので物理の話はしてないですし、量子力学で必要になることはほぼない話です。
ルベーグ積分が必要になるので、ほとんどがルベーグ測度とルベーグ積分の説明です。
最初に数学を無視した話をして、その後にやや数学よりの話をしますが、ほとんどがお話で数学的な証明はしてません。特に、抜け落ちている話がいくつもあることには注意してください。
知らなくてもなんとなく分かりそうな数学用語は説明なしで使っています

- 区間 (interval) が出てくるのその定義を与えておきます。開区間 (a,b) は $a<x<b$、閉区間 $[a,b]$ は $a\leq x\leq b$、半開区間 $[a,b),\ (a,b]$ は $a\leq x<b,\ a<x\leq b$ です。a,b は実数です。これらのどれかという意味で区間と言うことにします。区間の長さはどれも同じで $b-a$ です。
- d 次元実数空間を R^d と表記し、$d=1$ では R とします。
- 下限 inf が出てきます。簡単な例で言えば開区間 (x,y) での y が上限 (supremum) で、x が下限 (infimum) です。このとき y は最大値でなく、x は最小値ではないです。これに対して、閉区間では上限と最大値、下限と最小値は一致します。上限、下限は $\sup,\inf$ という記号が使われます。

L^2 空間の定義を与えるためにベクトル空間から始めます。実数または複素数を K とします (実数なら $\mathrm{K}=\mathrm{R}$、複素数なら $\mathrm{K}=\mathrm{C}$)。ある集合 V がありその元 $x,y\in V$ に対して、

- 和:$x+y\in V$
- 定数倍:$\alpha x\in V \quad (\alpha\in\mathrm{K})$

が定義されているとき、V はベクトル空間と呼ばれます。このとき

- 零ベクトル 0： $0x=0$
- x の逆ベクトル y： $x+y=0$

これらが存在するとされます。後は交換法則、結合法則、分配法則を持ちます。
次にノルム $\|\ \|$ を定義します。ノルムは $a\in\mathrm{K}$、$v,w\in V$ において

- $\|av\|=|a|\,\|v\|$
- $\|v+w\|\leq\|v\|+\|w\|$
- $\|v\|=0 \quad (v=0)$

を満たすものです。a は定数で $|a|$ は a の絶対値です。これらから、ノルムは

$$0=\|v-v\|\leq\|v\|+\|v\|=2\,\|v\|$$

となり、負にならないことが分かります。ノルムが定義されたベクトル空間はノルム空間 (normed space) と呼ばれます。ノルムはよくあるユークリッド空間の表記にすれば $|\boldsymbol{v}|=\sqrt{\boldsymbol{v}\cdot\boldsymbol{v}}$ です。
ノルム空間が完備であるときバナッハ空間 (Banach space) と呼ばれます。雑に言えば、コーシー列という数列がどこかに収束する空間です。少し細かく言えば、完備はコーシー列 (ある整数 $n_0>0$ があり $i,j>n_0$ で $\|v_i-v_j\|<\epsilon$ になる数列)$v_n\in V\ (n=1,2,\ldots)$ が、n の無限大で $v\in V$ と

$$\lim_{n\to\infty}\|v-v_n\|=0$$

となることです (コーシー列が同じ V で収束する)。
さらに内積を定義します。複素ベクトル空間において $v,w\in V$ を複素数 (v,w) にする写像として

- $(v,aw_1+bw_2)=a(v,w_1)+b(v,w_2)$

- $(av_1 + bv_2, w) = a^*(v_1, w) + b^*(v_2, w)$
- $(v, w) = (w, v)^*$
- $(v, v) \geq 0$
- $(v, v) = 0 \quad (v = 0)$

と定義されたものが内積で、内積が定義されていると内積空間と呼ばれます (ユークリッド空間では $(v, w) = \boldsymbol{v}\cdot\boldsymbol{w}$)。「$*$」は複素共役で、実ベクトル空間の内積 ($(v, w)$ は実数になる) では複素共役を外せばいいです。内積が定義されていれば、ノルムは $\sqrt{(v, v)}$ と与えられるので、内積空間はノルム空間です。完備な内積空間をヒルベルト空間と言い、内積空間はノルム空間なのでヒルベルト空間はバナッハ空間の特別な場合です。

ここで量子力学の要求を持ってきます。量子力学では波動関数 ψ によって、粒子がいる確率を

$$\int dx\ |\psi(x)|^2$$

によって与えます ($|\psi(x)|^2$ が確率密度)。確率が無限大になっては困るので

$$\int dx\ |\psi(x)|^2 < \infty \tag{1}$$

という制限が必要になります。このようなとき、関数 ψ を自乗可積分 (square integrable) と呼びます。量子力学の説明の中では、(1) を満たし、ノルムがこれから

$$\|\psi\| = \Big(\int dx\ |\psi(x)|^2\Big)^{1/2} \tag{2}$$

と与えられている関数 ψ の集まりを L^2 空間とし、内積を

$$(\psi, \phi) = \int dx \psi^\dagger(x)\phi(x) \tag{3}$$

と与えられるから L^2 空間はヒルベルト空間だと言います。しかし、これは L^2 空間の定義として不正確です。L^2 空間はバナッハ空間として定義されているので完備でなければいけないですが、この設定では完備になっていません。完備であるためには、コーシー列 ψ_n が

$$\lim_{n\to\infty}(\int dx\ |\psi_n(x) - \psi(x)|^2)^{1/2} = 0$$

となればいいですが、こうならない関数が存在します ($\lim \psi_n(x)$ が連続的でない関数がある)。

この状況をどうにかするためにルベーグ積分が必要になります。結果を言ってしまえば、バナッハ空間とするためには、まずノルム (2) においてリーマン積分で書かれている部分を、取りあえずルベーグ積分が何かは置いといて、ルベーグ積分

$$\int dx\ \Rightarrow\ \int_{\mathrm{R}} d\mu \tag{4}$$

に書き換えます。このときを $\mathcal{L}^2$ 空間とします。この段階ではノルムがまだ定義できていません。$\psi \neq 0$ のときに $\|\psi\| = 0$ となる場合があるからです。これを取り除いくことでノルムを (2) で定義でき、ノルム空間となります。そして、完備であることが示せるのでバナッハ空間となります。このバナッハ空間が L^2 空間です。L^2 空間はルベーグ空間 (Lebesgue space) とも呼ばれます（一般的には L^p 空間をルベーグ空間と呼ぶ。p は $|\psi(x)|^p$ から）。そして、内積を、内積 (3) に (4) の書き換えをしたものとして与えることでヒルベルト空間となります。

今行ったことは (4) に変えただけなので、見た目が変わっただけです。しかも、R^d での発散しないリーマン積分とルベーグ積分は一致します。このため、数学の事情を無視すれば、最初のリーマン積分による L^2 空間をヒルベルト空間と言っても物理の問題を扱うときに致命的なことは起こらないです。

ここからは、L^2 空間の数学的な定義を与えるために、実数の区間の場合に限定したルベーグ積分の簡単な話をしていきます。

まず、日常感覚的な範囲内に限定した面積の話をしていきます。面積と言った時の当たり前の性質から、面積を知りたい領域を細かい長方形で切っていけばその領域の面積に近づいていきます。この発想を逆側から行います。つまり、知りたい領域を小さく分割するのでなく、いくつかの長方形でお互いが重ならないように覆っていき、長方形を小さくして領域に対応させせようとします。これは、長方形を I_k とすれば、知りたい領域 A は $I = I_1 + \cdots + I_n$ の中に入るということです。

そうやって作られた面積を $S(A)$ とし、このときの $S(A)$ は I_k によって領域 A を最も上手く覆えるように選んだものとします（和が一番小さくなるように I_k の組み合わせを作った場合)。この $S(A)$ をルベーグ測度 (Legesgue measure) と呼び、A が辺の長さ x, y の長方形 L なら

$$S(L) = xy$$

となります。また、長方形 L を 2 つの長方形 L_1, L_2 に分離したとき、面積は当然

$$S(L_1 + L_2) = S(L_1) + S(L_2) = xy \tag{5}$$

と書けます。

今度はリーマン積分を見ます。リーマン積分は、関数 $f(x)$ の x を間隔 Δx_i で分割して、x_i での $f(x_i)$ との積 $f(x_i)\Delta x_i$ を足しあげたものです。$f(x)$ を分かりやすくするために、x が $a_i \leq x < b_i$ では定数 C_i ($f(x_i) = C_i$) となる関数で

$$0 < f(x_1) < \cdots < f(x_n)$$

となっているとします (階段状になっている)。このときのリーマン積分は

$$\int_a^b dx\ f(x) = f(x_1)(b_1 - a_1) + \cdots + f(x_n)(b_n - a_n)$$

となります ($a_1 = a$)。

今のように縦の分割 (x の分割) でなく、横に分割にしたと考えても同じ結果になることが分かると思います。つまり、$f(x)$ は $y_1, \ldots, y_n$ に分割されているとし、y_i での x は $b'_i - a'_i$、と読み替えます。今の $f(x)$ では、例えば $y_1 < C_1$ では x の区間は a_1 から b まで、$y_2 \geq C_1$ では a_2 から b までです。

ついでに、$b'_i - a'_i$ をルベーグ測度で書きます。ルベーグ測度は面積として与えましたが、1 次元では長さです (3 次元なら体積)。なので、区間の長さ $b'_i - a'_i$ をルベーグ測度 $S(E_i)$ と書くことにして (E_i は $b'_i - a'_i$ の区間を表す)

$$\int_E dx\ f(x) = \sum_{i=1}^{n} y_i S(E_i)$$

としたものをルベーグ積分 (Lebesgue integral) といいます。今の設定ではリーマン積分と同じ結果を出します。E は関数 f の定義域で、左辺の積分についてる E は E 上での積分という意味です。他にも R や範囲をより明確にして表記されたりします。

ルベーグ積分の特徴はルベーグ積分で書ける関数の極限にあります。話の流れから当たり前のようですが、ルベーグ測度 $S(E_i)$ での区間 E_i を与えられる関数 f の場合にルベーグ積分が与えられています。簡単に言えば、$x = f^{-1}(y)$ が E_i に入る関数です。このような関数を可測関数 (measurable function) と言い、可測関数の数列 f_n の極限に特徴があります。

$0 \leq f_i \leq f_{i+1}$ となる可測関数 $f_1, \ldots, f_n$ には、ルベーグ積分で

$$\int_E dx\ f_i(x) < \infty$$

となるなら、n が無限大の極限において

$$\lim_{n\to\infty} f_n(x) = f(x)$$

となる関数 f があり、f も可測関数であることが示せます (リーマン積分では極限の先 f でリーマン積分できなくてもいい)。そして、この性質から

$$\lim_{n\to\infty} \int_E dx\ f_n(x) = \int_E dx\ f(x) \tag{6}$$

という便利な性質を持つことが分かり、単調収束定理 (monotone convergence theory) と呼ばれます。そして、$0 \leq f_i \leq f_{i+1}$ としない可測関数に拡張したものを優収束定理 (dominated convergence theory) と言います ((6) と同じ形で、$|f_i(x)| \leq g(x)$ となるルベーグ積分が有限な関数 g がいるなら)。このような極限に対応する定理はリーマン積分にはないです。リーマン積分での極限について知ってる人なら分かると思いますが、リーマン積分で同じように極限を求めるのは非常に面倒です。

このようにルベーグ積分では極限に対する非常に便利な性質を持っています。これを利用することで、L^2 空間はバナッハ空間になるという Riesz-Fischer の定理を導けます (非常に面倒)。

今の話はかなり限定されている上に数学部分を全く考慮していません。なので、ルベーグ測度とルベーグ積分は上のように定義されたものと言ったら確実に文句を言われます。しかし、ルベーグ測度とルベーグ積分は今の話の流れで作られます。

次に、数学部分を多少考慮して同じ話を繰り返します。ここでも一般的でなく限定的な部分の説明だけです。また、証明はほとんど省いてます (面倒な証明しかないから)。

手順は、まず体積が満たすべき性質を持った関数を作り、それをルベーグ測度とします。次に関数 f を分割し、定義域にルベーグ測度をあてはめてルベーグ積分を作るというものです。

R^d の部分集合を A とします ($A \subset \mathrm{R}^d$)。まず、d 次元の閉区間 I を用意します。I は 1 次元なら線、2 次元なら長方形、3 次元なら直方体です。I の幅は

$$[a_i, b_i] \quad (i = 1, 2, \ldots, d)$$

によって与えます。そして A をこの閉区間 I をいくつも用意したものに覆われているとして

$$A \subset \bigcup_{k=1}^{\infty} I_k$$

とします。このとき閉区間の体積 (d 次元の体積) の和を

$$\sigma(S) = \sum_{k=1}^{\infty} v(I_k) \quad (v(I_k) = \prod_{i=1}^{d}(b_i^{(k)} - a_i^{(k)}),\ A \subset \bigcup_{k=1}^{\infty} I_k)$$

と定義します。A が I_k の和集合に含まれているとき、$S = \{I_k\}$ は A を覆う (cover) と言います。ルベーグ測度は、I_k によって A を覆ったときに最小になるものとしているので、下限 inf によって

$$\mu^*(A) = \inf_S \sigma(S)$$

とします。「$*$」は複素共役でなく、後の区別のためにつけてるだけです。inf の下についている S は A を覆える $S = \{I_k\}$ の選び方に対してという意味です。これを R^d でのルベーグ外測度 (Legesgue outer measure) と言います。区間の数が無限か有限かで性質の証明の仕方が変わりますが、ここではそんなこと一切気にせずに進めます。

また、何も注意せずに関数 μ^* を定義しましたが、これは集合を実数にする関数になっています。このような定義域を集合とする関数は、集合関数 (set function) と呼ばれます。なので、$\mu^*(A)$ と書いた時の A は集合です。

ルベーグ外測度は

$$\mu^*(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} \mu^*(A_i) \tag{7}$$

となっています。雑に示しておきます。定義から $\mu^*(A) \leq \sigma(S)$ なので

$$\mu^*(A_1 + A_2) \leq \sigma(S_1 + S_2)$$

$\sigma(S)$ は閉区間の和なので

$$\sigma(S_1 + S_2) \leq \sigma(S_1) + \sigma(S_2)$$

よって

$$\mu^*(A_1 + A_2) \leq \mu^*(A_1) + \mu^*(A_2)$$

となります。これは無限個でも成立していることが示せますが、省きます。大事な点は等号になっていないことです。μ^* を体積に対応するものだと考えるなら、(5) のように等号になっているべきです。なので、この段階では体積よりも大きな範囲のものを扱っていることになります。これを等号になるように制限します。

R^d の部分集合 A としていたのを閉区間 $I \subset \mathrm{R}^d$ としてみます。そうすると、定義を素直に受け取れば $\mu^*(I) = v(I)$ で (本来は証明が必要)、例えば R として $[a, b]$ なら $\mu^*([a, b]) = b - a$ です。そして、共通部分を持たない閉区間 I_k $(I_i \cap I_j = \emptyset,\ i \neq j)$ を考えれば、感覚的に

$$\mu^*(I_1 + I_2) = \mu^*(I_1) + \mu^*(I_2)$$

となり、実際に正しいことが証明されます。というわけで、μ^* は閉区間 I のとき体積の性質を再現しています。これは加算な I_n に対して

$$\mu^*(\bigcup_{n=1}^{\infty} I_n) = \sum_{n=1}^{\infty} \mu^*(I_n) \qquad (I_i \cap I_j = \emptyset,\ i \neq j) \tag{8}$$

として成立します。このように μ^* は一般的には (7) で、特別な場合として A が閉区間のとき (8) となります。

というわけで、欲しい性質を満たすものがルベーグ外測度に制限を加えることで作れます。(8) を満たすように定義域を取ったルベーグ外測度をルベーグ測度 (Legesgue measure) と言います。そのような集合を Σ とすれば、ルベーグ測度は $\mu = \mu^*|_\Sigma$ です。Σ（正確には Σ に含まれる集合）はルベーグ可測集合 (measurable set) と呼ばれます。可測集合と可測関数で同じ可測という単語を当ててますが、同じ意味になっているわけではないです。

ルベーグ測度は (8) の他に $\mu(\emptyset) = 0, \mu(E) \geq 0$ という性質を持っています。体積をこれらを満たすものとして抽象化したのが測度 (measure) です (体積が満たすべき性質を公理としたもの。下の補足参照)。その中で、R^d で測度の性質を満たすもの (R^d での可測集合 Σ による測度) がルベーグ測度です。

閉区間で見てきましたが、R^d の任意の区間 I はルベーグ可測集合です ($\mu(I) = \mu^*(I) = v(I)$)。最初の長方形の話では $\mu = \mu^*$ なので、ルベーグ外測度の定義でそのままルベーグ測度になっていたということです。また、R^d にはルベーグ測度になれない集合があることに注意してください。

ルベーグ測度が与えられたのでルベーグ積分に移ります。ここから $d = 1$ での R でのルベーグ測度しか出てきません (1 次元の積分だけ考える)。

先に定義関数を与えておきます。集合 S_i に対して、$x \in S_i$ であれば 1、$x \notin S_i$ であれば 0 となる関数 $\chi_{S_i}(x)$ を定義します。これは定義関数 (characteristic function) と呼ばれます。例えば、定義関数によって階段関数は定義されています。階段関数は、$\Theta(x) = \alpha$ $(1 \leq x < 2)$ でそれ以外では 0 とするような関数です。なので、階段関数は、階段関数に対する範囲の集合 S_i $(S_1 = [a_1, b_1), \ldots, S_n = [a_n, b_n))$ による定義関数によって

$$\Theta(x) = \alpha_1 \chi_{S_1}(x) + \cdots + \alpha_n \chi_{S_n}(x) \qquad (\alpha_i \in \mathrm{R})$$

と定義されます。

また、量子力学でも、箱 S の内部での積分を表すときに定義関数を使って

$$\int d^3x\ |\psi(x)|^2 \chi_S(x) = \int_S dx\ |\psi(x)|^2$$

と書かれることがあります。

ルベーグ積分では定義域でなく値域を分割するので、それを作ります。余計なことをしないですむので、$f(x)$ は 0 以上の実数とします。

$y = f(x)$ の y を分割して、$0 = y_0 < y_1 < y_2 < \ldots < y_n = \infty$ とします。定義域は $f^{-1}(y) = x$ で与えます。なので、$y_i \leq f(x) < y_{i+1}$ となる x の区間を E_i $(i = 0, 1, \ldots, n-1)$ とします。数学の記号で書けば

$$E_i = \{x \mid y_i \leq f(x) < y_{i+1}\}$$

となります。区間 E_i は重なっていないとします $(E_i \cap E_j = \emptyset,\ i \neq j)$。$E_i$ は区間なのでルベーグ可測集合です

次に $f(x)$ の分割を作ります。そのために、定義関数を使って $f(x)$ は

$$\sum_{i=0}^{n-1} y_i \chi_{E_i}(x)$$

と近似できるとします。$\chi_{E_i}(x)$ の定義から、例えば x_1 が E_1 の範囲に入っているとき $f(x_1) \simeq y_1$ となります。これは 1 つずらしたときでも $f(x)$ の近似になって

$$\sum_{i=0}^{n-1} y_{i+1} \chi_{E_i}(x)$$

このときは x_1 が E_1 の範囲に入っているとき $f(x_1) \simeq y_2$ となります。今は x は $y_i \leq f(x) < y_{i+1}$ に対してなので、これらは

$$\sum_{i=0}^{n-1} y_i \chi_{E_i}(x) \leq f(x) < \sum_{i=0}^{n-1} y_{i+1} \chi_{E_i}(x)$$

となります。

これを積分すれば和は無限大に取っていないので交換できて

$$\sum_{i=0}^{n-1} y_i \int_a^b dx\ \chi_{E_i}(x) \leq \int_a^b dx\ f(x) < \sum_{i=0}^{n-1} y_{i+1} \int_a^b dx\ \chi_{E_i}(x) \tag{9}$$

E_i は区間で、x がその区間内で $\chi_{E_i}(x) = 1$、区間の外で $\chi_{E_i}(x) = 0$ なので、積分はその区間の長さ $|E_i|$ です。そして、区間は重なっていないとしているので

$$\sum_{i=0}^{n-1} \int_a^b dx\ \chi_{E_i}(x) = b - a$$

となります。さらに、y_{i+1} と y_i の差の最大が ϵ になっているとすれば

$$\sum_{i=0}^{n-1} (y_{i+1} - y_i) \int_a^b dx\ \chi_{E_i}(x) \leq \epsilon(b - a)$$

となるので、(9) の差を小さく取ることができるのが分かります。この結果から

$$\sum_{i=0}^{n-1} y_i \int_a^b dx\ \chi_{E_i}(x) = \sum_{i=0}^{n-1} y_i |E_i| \simeq \int dx\ f(x)$$

と近似できます。

問題はないように進めてきましたが、$|E_i|$ が何か理解できないような状況になっていたら今のような話ができません。ここで出てくるのがルベーグ測度です。$|E_i|$ は区間の長さなので、1 次元の長さを抽象化したルベーグ測度に置き換えます。置き換えれば

$$\sum_{i=0}^{n-1} y_i\mu(E_i) \leq \int_a^b dx\ f(x) < \sum_{i=0}^{n-1} y_{i+1}\mu(E_i)$$

これを等号にして

$$\int_E dx\ f(x) = \sum_{i=0}^{n-1} y_i\mu(E_i) \tag{10}$$

としたのが $f(x) \geq 0$ に対するルベーグ積分です。f の定義域を E $(f:\ E \to \mathrm{R} \geq 0)$ で表しています。ルベーグ積分の表記は他にも

$$\int_X d\mu(x)\ f(x)\ ,\ \int_X d\mu\ f\ ,\ \int_X f$$

といったものがあります (μ でなく m や λ も使われます)。dx より $d\mu$ を使ったほうがリーマン積分と区別しやすいので、ここからはそれを使います。

ここで問題になるのはこれが可能な関数 f は何かですが、すでに $f(x)$ は定義関数で近似的に展開できるとしているので、結局

$$f(x) = \sum_{i=1}^{n} y_i\chi_{S_i}(x) \quad\ (y_i \in \mathrm{R})$$

となる関数なら可能です。S_i は $S_i \cap S_j = \emptyset$ $(i \neq j)$ です。このように書ける関数は単関数 (simple function) と呼ばれます。この展開は y_i が区別されていて、S_i が $S_i \cap S_j = \emptyset$ $(i \neq j)$ なら一意的に決まります (y_i への重なっていない区間の割り振り方は 1 つに決まる)。

というわけで、単関数 $f(x) \geq 0$ において、ルベーグ積分は f の定義域を E、ルベーグ可測集合を E_i として

$$\int_E d\mu(x)\ f(x) = \sum_{i=1}^{n} y_i\mu(E_i)$$

と定義されます。このように定義を与えると単関数は可測関数になっていますが、実際は証明が必要です。ちなみに、可測関数の定義は、区間 $[a,b]$ で関数 f が定義されているとき、$f(x) > c$ $(c \in \mathrm{R})$ となる $x \in [a,b]$ がルベーグ可測集合となるなら、f は可測関数と与えられています。なので、$x = f^{-1}(y)$ がルベーグ可測集合なら f は可測関数です。

可測関数の数列 $f_n(x)$ の極限が

$$\lim_{n\to\infty} f_n(x) = f(x)$$

と存在するので、単調収束定理と優収束定理が出てくるのは上で言ったのと同様です。

ルベーグ積分の分かりやすい例が階段関数です。上でした階段関数の定義から階段関数は単関数です。実際に (10) の定義に合わせて書けば

$$\int_A d\mu(x)\ \Theta(x) = \alpha_1(b_1 - a_1) + \cdots + \alpha_n(b_n - a_n)$$

となり、リーマン積分として実行しても同じ結果です。これは一般的な結果で、リーマン積分可能な関数 $f(x)$ の区間 E ($[a,b]$) に対して、リーマン積分とルベーグ積分は一致し

$$\int_a^b dx\ f(x) = \int_E d\mu(x)\ f(x)$$

となります。
気付いた人もいると思いますが、階段関数の例を使えばルベーグ積分の話を測度論なしで出来てしまいます (区間が明確に与えられているから)。そのような手順によるルベーグ積分の説明は MacNeille と Mikusinski によって与えられたようです (ここでしてきた説明はよくある手順との折衷)。ルベーグ積分の詳細が必要ないなら、この方法で十分だと思います。
ここでは 0 以上の実数の単関数 $f(x)$ に限定しましたが、0 以上の実数の可測関数、複素数の可測関数に対してもルベーグ積分は定義されています。
ルベーグ積分が何か分かったところで、L^2 空間の数学的な定義を与えます。必要になるのは考えてる空間 R^d、R^d 内のルベーグ可測集合 Σ、ルベーグ測度 μ です。これらがあるからルベーグ積分は定義できます。なので、$\mathrm{R}^d, \Sigma, \mu$ が与えられているとして、$\mathcal{L}^2$ 空間を

$$\int_{\mathrm{R}^d} d\mu\ |f|^2 < \infty$$

となる関数 f の集まりとします。数学記号で書けば

$$\mathcal{L}^2 = \{f \mid \int_{\mathrm{R}^d} d\mu\ |f|^2 < \infty\}$$

となります。しかし、(10) を見れば分かるように

$$\|f\| = (\int_{\mathrm{R}^d} d\mu\ |f|^2)^{1/2} \tag{11}$$

をノルムにしようとしても、$\mu = 0$ が可能なために $\|f\| = 0$ で $f = 0$ を意味しないので、ノルムになりません。
これを回避するために、関数 f, g が $\mu(x) = 0$ ($x \in \mathrm{R}^d$) となる範囲を除いて $f(x) = g(x)$ のとき f, g は等しいとする定義を加えます。この意味で $f = g$ almost everywhere と表記します。逆に言えば、$f(x) \neq g(x)$ となる x において $\mu(x) = 0$ ということです。つまり

$$\|f - g\| = 0\ \Leftrightarrow\ f = g \quad \text{almost everywhere}$$

とします。almost everywhere として、$f = 0$ のとき $\|f\| = 0$ となります。almost everywhere は a.e. と略されます。余計なことを加えたくなかったので、上の話では almost everywhere であることを無視している部分があります。
というわけで、$\mathcal{L}^2$ 空間に $f = g$ almost everywhere を加えたものが L^2 空間となります (ノルムは (11))。これに内積を加えることでヒルベルト空間となります。

・補足
測度論で出てくる数学の定義を並べておきます。数学用語の説明は省いています。
いきなり定義を書いても理由がわかりづらいので前振りをしておきます。まず、体積であるための条件を与えます。条件は通常の体積で成立しているものです。条件の 1 つは、体積は 0 以上ということです。これを集合の言葉にします。体積が 0 のときは対象となるものがないことなので、単純に空集合 $\emptyset$ を使えばいいです。0 以上であることは、ある集合 S を実数にする集合関数 μ によって与えます (記号が同じですがルベーグ測度としていません)。つまり、

$$\text{(i)}: \mu(\emptyset) = 0\ ,\ \mu(S) \geq 0$$

とします。

　もう1つの条件は、重なっていない2つの領域の体積は、同時に体積を求めた場合と、個別に求めた場合で一致するというものです(例えば、ある直方体の体積と、それを水平に切って分離させた2つの直方体の体積の和は一致)。これは、二つの集合 S_1, S_2 があり、互いに素のとき $(S_1 \cap S_2 = \emptyset)$

$$\mu(S_1 + S_2) = \mu(S_1) + \mu(S_2)$$

ということです。なので、加算な S_i に対して

$$\text{(ii)}:\mu(\bigcup_{i=1}^{\infty} S_i) = \sum_{i=1}^{\infty} \mu(S_i) \quad (S_i \cap S_j = \emptyset,\ i \neq j)$$

これら2つの条件を満たす μ の定義を与えます。

　ここで必要になるのが (i),(ii) となる集合が何かです。話を省くために、ルベーグ測度は区間によって作れていたのを利用します。全ての区間の集合族 (集合の集合) を $\mathcal{I}$ とし、これには点と空集合も含むようにします。そして、区間の全ての和集合の集合族を $\mathcal{B}$ とします。$\mathcal{B}$ は

$$I_1 \cup I_2 \cup \cdots \cup I_n \quad (I_k \in \mathcal{I})$$

と与えられます。$\mathcal{B}$ は区間の和集合なので、$S, T \in \mathcal{B}$ のとき

$$S = I_1 \cup I_2 \cup \cdots \cup I_m\ ,\ T = J_1 \cup J_2 \cup \cdots \cup J_n \quad (I_1, \ldots, I_m, J_1, \ldots, J_n \in \mathcal{I})$$

なので、

$$S \cup T = I_1 \cup \cdots \cup I_m \cup J_1 \cup \cdots \cup J_n$$

よって $\mathcal{B}$ にいるので、$S \cup T \in \mathcal{B}$ です。

　次に補集合を見てみます（補集合は添え字 c で表します)。一般的にみるのは面倒なので、$S = I_1 = [0, 1]$ とします。このとき I_1 の補集合 I_1^c は

$$I_1^c = (-\infty, 0) \cup (1, \infty)$$

と書けます。つまり、区間の和集合です。なので、$S = I_1$ の補集合は $\mathcal{B}$ にいます。これはそのまま一般的に言えて、$S^c \in \mathcal{B}$ です。というわけで、区間を考えるとき、集合の関係として

$$S \cup T \in \mathcal{B}\ ,\ S^c \in \mathcal{B} \quad (S, T \in \mathcal{B})$$

が出てきます。

　ここで取り出した関係は、体積が与えられる領域において、領域同士をくっつけても体積が与えられる (S, T の体積が与えられるなら $S \cup T$ も)、領域内の一部の体積が与えられるなら残りの部分の体積も与えられる (S の体積が与えられるなら S^c も)、と言っているようなものです。このような関係を持つ区間によって、(i),(ii) を満たすルベーグ測度は作られていました。なので、この関係を満たす集合を作れば μ の定義に使えると考えます。

　この関係を取り出して定義を与えます。集合 X の全ての部分集合による集合族 $\mathcal{A}$ があり、$\mathcal{A}$ の元 ($\mathcal{A}$ の元は集合) と空集合 $\emptyset$ に対して

- $A \cup B \in \mathcal{A} \quad (A, B \in \mathcal{A})$
- $\emptyset \in \mathcal{A}$

そして、補集合 A^c に対して

- $A^c \in A \quad (A \in \mathcal{A})$

であるとき、$\mathcal{A}$ はブール代数 (Boolean algebra) や代数と呼ばれます。
これに加算な A_i の和集合が

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A} \quad (A_n \in \mathcal{A})$$

となることを加えたブール代数を σ 代数と呼びます。例えば X の全ての部分集合による集合族 $\mathcal{P}(X)$(冪集合, power set) は定義から明らかに σ 代数です。

区間からの類推のみなので結論に行くには飛びがありますが (Carathéodory 条件の話が抜けてる)、σ 代数によって μ に (i),(ii) を持たせられます。つまり、集合から 0 以上の実数を作る写像 μ が、$\mathcal{A}$ を σ 代数として

- $\mu(\emptyset) = 0$
- $\mu(A) \geq 0 \quad (A \in \mathcal{A})$
- $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n) \quad (A_n \in \mathcal{A}\ ,\ A_m \cap A_n = \emptyset\ (n \neq m))$
- $\emptyset \in \mathcal{A}$
- $A^c \in \mathcal{A} \quad (A \in \mathcal{A})$
- $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A} \quad (A_n \in \mathcal{A})$

を満たすとします。この μ を測度 (measure) と呼びます。$\mu(A)$ は無限大にも取れるとし (μ は $\mathcal{A}$ から $\mathrm{R}_{\geq 0} \cup \{\infty\}$ への写像)、$\mu(A) < \infty$ なら有限の測度と呼ばれます。

集合 X と X の部分集合の σ 代数 $\mathcal{A}$ による組 $(X, \mathcal{A})$ は可測空間 (measureable space)、$\mathcal{A}$ に含まれる集合 A は可測集合 (measureable set) と呼ばれます。可測空間 $(X, \mathcal{A})$ は簡単に言えば、σ 代数 $\mathcal{A}$ による構造を持つ集合 X によるものということです。A は σ 代数の性質を持つので、A が可測集合ならその補集合は可測、可算な可測集合 A_i の和集合 $\cup_{i=1}^{\infty} A_i (i = 1, 2, \ldots, \infty)$ は可測です。可測集合 A_i があり、$\mu(A_i)$ が有限で、集合 A が可算の和集合 $\cup_{i=1}^{\infty} A_i$ であるなら $\mu(A)$ は σ 有限 (σ finite) と呼ばれます。

可測空間 $(X, \mathcal{A})$ に測度 μ を加えた $(X, \mathcal{A}, \mu)$ を測度空間 (measure space) と言います。つまり、$(X, \mathcal{A}, \mu)$ は集合 X、X での σ 代数 $\mathcal{A}$、$\mathcal{A}$ での測度 μ によって与えられています。$\mu(X) = 1$ では確率空間と呼ばれます。

測度空間 $(X, \mathcal{A}, \mu)$ が与えられているとき、ルベーグ積分は作れて

$$\int_X f d\mu$$

となります。f は $\mathcal{A}$ での可測関数です。

ついでに関連する単語を並べておきます。

集合族$\mathcal{U}$ を実数や複素数にする関数 F を集合関数 (set function) と言います。可算な $U_1, U_2, \ldots \in \mathcal{U}$ が $U_i \cap U_k = \emptyset\ (i \neq k)$ のとき

$$F(U_1 \cup U_2 \cup \cdots) = F(U_1) + F(U_2) + \cdots$$

であるなら、$F(U)$ は可算加法的 (countably additive) や σ 加法的と言います。

X の部分集合による集合族 $\mathcal{F}$ があるとします。このとき、$\mathcal{F}$ を含む最小 (smallest) の σ 代数 $\mathcal{A}(\mathcal{F})$ が一意的に存在します。$\mathcal{A}(\mathcal{F})$ は $\mathcal{F}$ から生成される σ 代数 (σ algebra generated by $\mathcal{F}$) と呼ばれます。最小は、集合族 $\mathcal{X}$ に含まれる集合 U が $\mathcal{X}$ での他の任意の集合に含まれているときに使われます。なので、$\mathcal{F}$ から生成される σ 代数 $\mathcal{A}(\mathcal{F})$ は $\mathcal{F}$ を含む全ての σ 代数の共通部分です。

生成される σ 代数としてボレル (Borel)σ 代数というのがあります。ボレル σ 代数は、ベクトル空間 (距離空間) での集合 X での開集合の集合によって生成される σ 代数 $\mathcal{B}(X)$ です。ボレル σ 代数に含まれる集合はボレル集合と呼ばれます (ボレル σ 代数はボレル集合の集合なのでボレル集合族とも呼ばれます)。例えば R の開区間から生成される σ 代数はボレル σ 代数です (このボレル集合はルベーグ可測集合に対応する)。