Существование такого интеграла обеспечивается тем, что в произведении $f(x)e(x)$ один множитель суммируем, тогда как другой измерим и ограничен. Вообще говоря, произведение $f(x)e(x)$ может быть измеримо без того, чтобы функция $f(x)$ была измерима на всем ее протяжении или чтобы было измеримо множество e; что касается этого последнего, то достаточно, чтобы была измерима та его часть, на которой $f(x) \neq 0$. Ясно, что если $f(x)$ суммируема на множестве e, то она суммируема и на любом его измеримом подмножестве.

Так как мера множества e есть не что иное, как интеграл от функции, равной 1 на e, то естественно поставить вопрос, присущи ли интегралу, который мы только что определили, установленные выше свойства меры. Ответ на этот вопрос, как легко видеть, может быть сформулирован следующим образом:

Если e_1, e_2, ... — непересекающиеся измеримые множества и если функция $f(x)$ измерима на $e_1 \cup e_2 \cup \ldots$, то

$$F(e_1 \cup e_2 \cup \ldots) = F(e_1) + F(e_2) + \ldots,$$

т. е. $F(e)$ представляет собой счетно-аддитивную функцию множества.

Свойство *абсолютной непрерывности* неопределенного интеграла выступает здесь в следующей обобщенной форме:

Если измеримое множество e изменяется так, что его мера стремится к нулю, то $F(e)$ также стремится к нулю; другими словами, для каждого $\varepsilon > 0$ существует такое $\delta > 0$, что $|F(e)| < \varepsilon$, коль скоро $m(e) < \delta$.

Доказать это можно так же, как тогда, когда e образовано непересекающимися интервалами, представив $f(x)$ в виде суммы двух функций, из которых одна ограничена, а абсолютная величина другой имеет сколь угодно малый интеграл.

§ 3. ПРОСТРАНСТВО L^2 И ЛИНЕЙНЫЕ ФУНКЦИОНАЛЫ В НЕМ. ПРОСТРАНСТВА L^p

28. Пространство L^2. Сходимость в среднем. Теорема Рисса — Фишера. Рассмотрим измеримые функции $f(x)$, определенные на некотором множестве e конечной или бесконечной меры. Имея в виду дальнейшие приложения, допустим, что $f(x)$ могут принимать комплексные значения. Наконец, мы предположим, что для каждой f суммируем квадрат ее модуля $|f|^2$; в силу неравенства Шварца, сами f суммируемы на всех подмножествах множества e, имеющих конечную меру. Из того же неравенства следует, что произведение любых двух таких функций также суммируемо.

Обозначим

$$(f, g)=\int_e f(x)\overline{g(x)}\,dx, \qquad \|f\|=\sqrt{(f, f)}=\left(\int_e |f(x)|^2 dx\right)^{1/2}$$

и назовем первое из этих выражений *скалярным произведением* f и g, а второе — *нормой* функции f; последняя всегда положительна, за исключением того случая, когда $f(x)=0$ почти всюду. Согласно неравенству Шварца,

$$|(f, g)|\leqslant\|f\|\|g\|;$$

далее, из неравенства Минковского вытекает, что если f и g принадлежат рассматриваемому классу, то ему принадлежит и $f+g$, причем

$$\|f+g\|\leqslant\|f\|+\|g\|.$$

Заметив, что вместе с любыми двумя функциями их разность также принадлежит рассматриваемому классу, и взяв f_1-f_2 вместо f и f_2-f_3 вместо g, получим „неравенство треугольника"

$$\|f_1-f_3\|\leqslant\|f_1-f_2\|+\|f_2-f_3\|.$$

Оно напоминает неравенство между длинами сторон треугольника, если „расстояние" от f до g определено как $\|f-g\|$. При этом естественно *не различать две функции, расстояние между которыми равно нулю*, т. е. функции, совпадающие почти всюду; это соглашение будет соблюдаться всюду в дальнейшем.

Очевидно также, что одновременно с f и g нашему классу принадлежат λf и λg и при этом

$$(\lambda f, g)=\lambda(f, g), \quad (f, \lambda g)=\bar\lambda(f, g), \quad \|\lambda f\|=|\lambda|\|f\|;$$

отметим еще соотношение

$$\overline{(g, f)}=(f, g).$$

Все это делает очевидной аналогию между выделенным классом функций, называемым *(комплексным) функциональным пространством L^2*, и обычным (комплексным) векторным пространством[1]. Основное различие между тем и другим состоит в том, что L^2 не является конечномерным пространством, т. е. его элементы не исчерпываются линейными комбинациями какого бы то ни было конечного числа их. Именно это различие определяет собой важность некоторых понятий, теряющих интерес в случае конечномерных пространств.

Первым из таких понятий является *сходимость в среднем*, иначе называемая *сильной сходимостью*. Мы говорим, что последовательность $\{f_n\}$ сходится в среднем к f, если $\|f-f_n\|\to 0$. (В дальнейшем при изложении абстрактной теории гильбертова

---

[1]) *Действительное* L^2 получится, если все f принимают только действительные значения; при этом $(f, g)=(g, f)$.

пространства мы будем это же соотношение записывать в виде $f_n \longrightarrow f$; теперь же, пока речь идет о функциях, такая запись могла бы привести к недоразумениям.)

К этому понятию относится следующая полезнейшая теорема, аналогичная классическому критерию сходимости Коши:

Теорема Рисса—Фишера[1]. *Для того чтобы существовал элемент f, к которому заданная последовательность $\{f_n\}$ сходится в среднем, необходимо и достаточно условие: $\|f_m - f_n\| \longrightarrow 0$ при m, $n \longrightarrow \infty$.*

Высказанное условие необходимо; в самом деле, в силу неравенства треугольника,

$$\|f_n - f_m\| \leqslant \|f - f_n\| + \|f - f_m\| \longrightarrow 0.$$

Для доказательства достаточности выберем номера $m_1 < m_2 < \ldots$ таким образом, чтобы при $n > m_k$ выполнялись неравенства

$$\|f_n - f_{m_k}\| < 2^{-k}.$$

Тогда, в частности, будет $\|f_{m_{k+1}} - f_{m_k}\| < 2^{-k}$ и так как, согласно неравенству Шварца с $g(x) = 1$,

$$\int_{e'} |f_{m_{k+1}}(x) - f_{m_k}(x)|\, dx \leqslant \sqrt{m(e')}\, \|f_{m_{k+1}} - f_{m_k}\| \leqslant \sqrt{m(e')} \cdot 2^{-k}$$

для любого подмножества e' множества e, мера $m(e')$ которого конечна, то ряд

$$\sum_1^\infty \int_{e'} |f_{m_{k+1}}(x) - f_{m_k}(x)|\, dx$$

будет сходиться. Отсюда, в силу теоремы Б. Леви (п. 18), ряд

$$\sum_1^\infty [f_{m_{k+1}}(x) - f_{m_k}(x)]$$

почти всюду сходится абсолютно, и, следовательно, $f_{m_k}(x)$ почти всюду стремится к некоторому пределу $f(x)$. Далее, неравенства

$$\|f_{m_k}\| \leqslant \|f_{m_1}\| + \|f_{m_k} - f_{m_1}\| \leqslant \|f_{m_1}\| + \frac{1}{2}$$

показывают, что нормы $\|f_{m_k}\|$, а следовательно, и интегралы от $|f_{m_k}|^2$ ограничены; в силу леммы Фату (п. 20), $f(x)$ принадлежит пространству L^2. Применив ту же лемму к последовательности $\{f_{m_k} - f_n\}$, где n фиксировано и $n > m_r$, и заметив, что при $k > r$

$$\|f_{m_k} - f_n\| \leqslant \|f_{m_k} - f_{m_r}\| + \|f_n - f_{m_r}\| < 2^{-r+1},$$

---

[1] Рисс [1], [2]; Фишер [1].

получим

$$\|f-f_n\| \leqslant 2^{-r+1}.$$

При $n \to \infty$ целое положительное число r может быть взято сколь угодно большим; следовательно,

$$\|f-f_n\| \to 0,$$

что и требовалось доказать. Заметим, наконец, что функция f — предел в среднем последовательности $\{f_n\}$ — определена с точностью до множества меры нуль; в самом деле, если f^* служит пределом в среднем той же последовательности $\{f_n\}$, то

$$\|f-f^*\| \leqslant \|f-f_n\| + \|f_n-f^*\| \to 0,$$

т. е. $\|f-f^*\|=0$.

29. Слабая сходимость. Если последовательность $\{f_n\}$ сходится в среднем к f, то

$$(f_n, g) \to (f, g), \tag{7}$$

каков бы ни был элемент g пространства L^2; это вытекает непосредственно из оценки

$$|(f, g)-(f_n, g)| = |(f-f_n, g)| \leqslant \|f-f_n\| \|g\| \to 0.$$

Обратное неверно: если (7) справедливо для любого g, то отсюда не следует сходимость в среднем, как показывает пример $f_n(x) = \sin nx$, $0 \leqslant x \leqslant \pi$, заимствованный из теории рядов Фурье.

Говорят, что *последовательность* $\{f_n\}$ *сходится слабо к* f, если для всех g выполняется соотношение (7) или эквивалентное ему соотношение $(g, f_n) \to (g, f)$. В приведенном выше примере последовательность $\{\sin nx\}$ сходится слабо к 0 и в то же время не сходится в среднем ни к какой функции, так как при $m \neq n$

$$\|\sin mx - \sin nx\|^2 = \pi.$$

Одно из существенных различий между обоими видами сходимости состоит в том, что в случае сходимости в среднем

$$\|f_n\| \to \|f\|, \tag{8}$$

так как

$$\|f_n\| \leqslant \|f_n-f\| + \|f\|, \quad \|f\| \leqslant \|f-f_n\| + \|f_n\|;$$

в случае слабой сходимости соотношение (8), вообще говоря, не выполняется; так, в указанном выше примере $f_n = \sin nx$, $0 \leqslant x \leqslant \pi$, имеем $\|f_n\| = \sqrt{\frac{\pi}{2}}$, тогда как $\|f\|=0$.

Впрочем, слабая сходимость при выполнении соотношения (8) влечет за собой сходимость в среднем, так как

$$\|f-f_n\|^2 = \|f\|^2 - (f, f_n) - (f_n, f) + \|f_n\|^2 \to$$
$$\to \|f\|^2 - (f, f) - (f, f) + \|f\|^2 = 0.$$

Относительно слабой сходимости норма обнаруживает нечто вроде полунепрерывности, а именно,

$$\|f\| \leqslant \liminf \|f_n\|;$$

этому условию эквивалентно следующее: если $\|f_n\| \leqslant C$, то $\|f\| \leqslant C$. Это следует из соотношений

$$C\|f\| \geqslant |(f_n, f)| \to (f, f) = \|f\|^2.$$

30. Линейные функционалы. Понятие слабой сходимости станет более отчетливым, если ввести понятие линейного функционала. В самом деле, скалярное произведение (g, f) при фиксированном f и переменном g представляет собой линейную функцию от g или, согласно общепринятой терминологии, линейный функционал.

Вообще *линейным функционалом* в L^2 называется операция A, ставящая в соответствие любому элементу g пространства L^2 число Ag, если она

1) *аддитивна*: $A(g_1 + g_2) = Ag_1 + Ag_2$;
2) *однородна*: $A(cg) = cAg$, где c — любой числовой множитель;
3) *ограничена*: существует число M, такое, что для любого g

$$|Ag| \leqslant M\|g\|.$$

Наименьшее из этих чисел M обозначается M_A или $\|A\|$ и называется *нормой* линейного функционала.

Мы видим, что функционал $Ag = (g, f)$ отвечает всем этим требованиям, и для него $M_A = \|f\|$. Значительно более важна обратная

Теорема[1]). *Всякий линейный функционал Ag в пространстве L^2 может быть представлен в виде*

$$Ag = (g, f),$$

причем „производящая функция" f однозначно определяется заданием функционала A.

Доказательство основывается на формуле

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2, \tag{9}$$

справедливой для любых элементов u и v пространства L^2. Устанавливается она непосредственным подсчетом, и геометрическая ее интерпретация состоит, очевидно, в том, что в параллелограмме сумма квадратов диагоналей равна сумме квадратов сторон.

Рассмотрим линейный функционал A; верхняя грань $|Ag|$ на „единичной сфере" $\|g\| = 1$ равна M_A, поэтому существует последовательность $\{g_n\}$, такая, что

$$\|g_n\| = 1, \quad |Ag_n| \to M_A.$$

---

[1]) Эту теорему доказали независимо друг от друга Фреше и Рисс и опубликовали ее одновременно в 144 томе Comptes Rendus (1907). Приведенное здесь доказательство принадлежит Риссу [16].

Снабдив, если нужно, функции g_n некоторыми числовыми множителями, равными по модулю 1, мы можем добиться того, что все Ag_n станут действительными и неотрицательными, так что Ag_n будет стремиться к M_A. Положив в (9) $u=g_m$, $v=g_n$ и заметив, что

$$Ag_m+Ag_n=A(g_m+g_n)\leqslant M_A\|g_m+g_n\|,$$

получим

$$\|g_m-g_n\|^2=4-\|g_m+g_n\|^2\leqslant$$
$$\leqslant 4-\frac{1}{M_A^2}(Ag_m+Ag_n)^2\to 4-\frac{1}{M_A^2}\cdot 4M_A^2=0.$$

В силу теоремы Рисса — Фишера, последовательиость $\{g_n\}$ сходится в среднем к некоторому элементу g^*. Далее, $\|g^*\|=1$, и так как

$$|Ag^*-Ag_n|=|A(g^*-g_n)|\leqslant M_A\|g^*-g_n\|\to 0,$$

т. е. $Ag_n\to Ag^*$, то $Ag^*=M_A$. Итак, линейный функционал Ag на единичной сфере достигает при $g=g^*$ своей верхней грани M_A.

Покажем теперь, что

$$Ag=(g, f), \tag{10}$$

где $f=M_Ag^*$. Это равенство очевидно для $g=g^*$. Далее, g^* ортогонален всем тем g, для которых $Ag=0$; это означает, что

$$(g, g^*)=(g^*, g)=0.$$

В самом деле, для таких g имеем

$$M_A^2=(Ag^*)^2=(A(g^*-\lambda g))^2\leqslant M_A^2\|g^*-\lambda g\|^2=$$
$$=M_A^2(1-\lambda(g, g^*)-\bar{\lambda}(g^*, g)+\lambda\bar{\lambda}(g, g)),$$

откуда следует, что

$$-\bar{\lambda}(g^*, g)-\lambda(g, g^*)+\lambda\bar{\lambda}(g, g)\geqslant 0;$$

положив

$$\lambda=\frac{(g^*, g)}{(g, g)},\quad \bar{\lambda}=\frac{(g, g^*)}{(g, g)},$$

получим

$$-|(g, g^*)|^2\geqslant 0.$$

Таким образом, $(g, g^*)=0$ и, следовательно, $(g, f)=M_A(g, g^*)=0$. Итак, формула (10) установлена для $g=g^*$ и для всех g_0, обращающих функционал Ag в нуль. Так как любой элемент g может быть представлен в виде $g=g_0+\mu g^*$, если взять

$$\mu=\frac{Ag}{M_A}=\frac{Ag}{Ag^*},$$

то формула (10) оказывается справедливой для произвольного g, что и требовалось доказать.

31. Последовательности линейных функционалов. Теорема Осгуда. Возвращаясь к понятию слабой сходимости, заметим, что для функционалов слабая сходимость f_n к f означает, что последовательность функционалов $A_n g = (g, f_n)$ сходится всюду в L^2 к функционалу $Ag = (g, f)$. С этой точки зрения результат, который мы собираемся установить, оказывается связанным с началами теории функций действительного переменного. Сначала будет сформулирована

Теорема Осгуда[1]). *Если последовательность непрерывных функций $\{f_n(t)\}$, заданных в интервале (a, b), ограничена при любом значении t внутри (a, b), то в некотором интервале, заключенном в (a, b), эта последовательность равномерно ограничена.*

Допустим, что для заданной последовательности $\{f_n(t)\}$ теорема неверна. Тогда существуют номер n_1 и точка t_1 внутри интервала (a, b), такие, что $|f_{n_1}(t_1)| > 1$; в силу непрерывности f_{n_1}, найдется интервал (a_1, b_1), окружающий t_1, в котором $|f_{n_1}(t)| > 1$. Точно так же существуют номер $n_2 > n_1$ и интервал (a_2, b_2), заключенный внутри (a_1, b_1), такие, что $|f_{n_2}(t)| > 2$ в (a_2, b_2). Продолжая рассуждать таким образом, мы получим последовательность номеров $n_1 < n_2 < \ldots$ и последовательность интервалов (a_k, b_k), каждый из которых заключен в предыдущих, такие, что $|f_{n_k}(t)| > k$ в (a_k, b_k). Так как, очевидно, можно предположить, что $a_k < a_{k+1}$, $b_k > b_{k+1}$ и $b_k - a_k \to 0$, то найдется точка t_0, принадлежащая одновременно всем интервалам (a_k, b_k), в которой, в противоречии с предположением, $|f_{n_k}(t_0)| > k$ для всех k.

Мы привели здесь эту теорему и ее доказательство потому, что, почти не изменяя хода рассуждений, мы можем прийти к важным результатам, касающимся последовательностей линейных функционалов и даже более общих операций, притом не только в пространстве L^2 или в пространствах L^p, которыми мы займемся ниже, но в самых разнообразных пространствах, как функциональных, так и абстрактных. Для определенности мы ограничимся здесь рассмотрением последовательностей линейных функционалов в пространстве L^2. Результат, который мы собираемся установить, состоит в следующем:

Теорема[2]). *Если последовательность линейных функционалов $\{A_n g\}$ сходится или хотя бы ограничена при любом g, то она равномерно ограничена на единичном шаре, т. е. ограничена последовательность норм $\{M_{A_n}\}$.*

Первая часть доказательства повторяет предыдущее рассуждение, только надо взять A_n вместо f_n, элементы g пространства L^2 вместо значений t и какой-нибудь шар, например единичный,

---

[1]) Осгуд [1] (стр. 159—164).

[2]) Банах и Штейнгауз [1], Банах [3] (стр. 80); см. также Лебег [5] (стр. 70).

вместо интервала (a, b). Интервалам (a_k, b_k) в этом доказательстве будут соответствовать шары с центрами в g_{n_k} и радиусами $\rho_k < \frac{1}{k}$, настолько малыми, чтобы каждый шар был заключен внутри предыдущего. Так как при $k < l$

$$\|g_{n_l} - g_{n_k}\| < \frac{1}{k},$$

то последовательность $\{g_{n_k}\}$ сходится в среднем к некоторому элементу g_0, лежащему внутри всех выделенных шаров, и

$$|A_{n_k} g_0| > k \to \infty,$$

в противоречии с предположением. Итак, в некотором шаре с центром в g^* и радиусом ρ, т. е. на множестве элементов вида $g^* + \rho g$, где $\|g\| \leqslant 1$, линейные функционалы A_n равномерно ограничены; это означает, что

$$|A_n(g^* + \rho g)| \leqslant C_1,$$

где C_1 — некоторая постоянная, а g — переменный элемент единичного шара. Для этих последних будем тогда иметь

$$|A_n g| = \frac{1}{\rho} |A_n(g^* + \rho g) - A_n g^*| \leqslant \frac{1}{\rho} |A_n(g^* + \rho g)| + \frac{1}{\rho} |A_n g^*| \leqslant$$
$$\leqslant \frac{2}{\rho} C_1 = C_2,$$

и, таким образом, теорема доказана.

В терминах слабой сходимости эту теорему можно высказать следующим образом: *всякая слабо сходящаяся последовательность $\{f_n\}$ ограничена,*

$$\|f_n\| \leqslant C.$$

32. Сепарабельность пространства L^2. Теорема выбора. Доказанная выше теорема в известном смысле обратима, а именно, справедлив следующий аналог теоремы Больцано — Вейерштрасса:

Теорема выбора. *Из любой ограниченной последовательности $\{f_n\}$, $\|f_n\| \leqslant C$, элементов пространства L^2 можно выделить слабо сходящуюся подпоследовательность.*

Доказательство основывается на том, что пространство L^2 *сепарабельно*, т. е. оно содержит такое счетное подмножество элементов $\{g_n\}$, что всякий элемент f пространства L^2 может быть сколь угодно точно приближен в среднем этими g_n или их линейными комбинациями $l = c_1 g_1 + \ldots + c_m g_m$. Докажем сначала это предложение. Какова бы ни была область определения функции f, она всегда заключена в интервале $(-\infty, \infty)$; далее, если отдельно действительная и мнимая части функции f приближены функциями l_1 и l_2 с точностью до ε, то f приближена функцией $l_1 + il_2$

с точностью до 2ε. Поэтому достаточно рассмотреть „действительное" L^2 с элементами $f(x)$, заданными на всей оси x. Для каждой такой $f(x)$ пусть $f_n(x)$ означает функцию, при $|x| \leqslant n$ равную $f(x)$, если $|f(x)| \leqslant n$, и равную n или $-n$, соответственно если $f(x) > n$ или $f(x) < -n$, а при $|x| > n$ равную нулю. Тогда $f_n \to f$ почти всюду и $(f - f_n)^2 \leqslant f^2$. Таким образом, получаем, что $\|f - f_n\| \to 0$, т. е. f может быть сколь угодно точно приближена в указанном смысле ограниченной функцией, равной нулю вне некоторого конечного интервала (a, b). В то же время любая функция последнего типа представляет собой предел почти всюду некоторой последовательности ступенчатых функций $\{\varphi_n\}$, заданных и равномерно ограниченных на (a, b). Согласно „малой" теореме Лебега (п. 19), эта последовательность будет к своему пределу сходиться в среднем. Наконец, всякая ступенчатая функция может быть приближена ступенчатыми же функциями, но имеющими разрывы лишь в рациональных точках. Итак, переходя последовательно от одного класса к другому, каждый раз более узкому, и пользуясь неравенством треугольника, мы приходим к заключению, что произвольный элемент f может быть сколь угодно точно приближен в среднем ступенчатыми функциями с разрывами в рациональных точках, т. е. линейными комбинациями характеристических функций интервалов с рациональными концами. Так как интервалы такого рода образуют счетное множество, то сепарабельность пространства L^2 доказана.

Переходим к доказательству теоремы выбора. Последовательность $\{f_n\}$, по предположению, ограничена, т. е. $\|f_n\| \leqslant C$, следовательно, $|(g, f_n)| \leqslant C\|g\|$; это означает, что, каков бы ни был элемент g, скалярные произведения (g, f_n) образуют ограниченную числовую последовательность. Пользуясь теоремой Больцано — Вейерштрасса и классическим диагональным методом, мы выделим подпоследовательность $\{f_{n_k}\}$, для которой сходятся последовательности скалярных произведений (g, f_{n_k}) для любого элемента g из указанного выше счетного множества, а также для любой линейной комбинации l таких элементов. Положив

$$A_k g = (g, f_{n_k}),$$

мы скажем, что последовательность линейных функционалов $\{A_k\}$ ограничена по норме, т. е. $M_{A_k} \leqslant C$, и сходится на некотором множестве $\{l\}$, всюду плотном в L^2 в смысле сходимости в среднем. Соотношения же

$$\begin{aligned}|A_n g - A_m g| \leqslant |A_n g - A_n l| + |A_n l - A_m l| + |A_m l - A_m g| \leqslant \\ \leqslant 2C\|g - l\| + |A_n l - A_m l|\end{aligned}$$

обеспечивают сходимость $\{A_n g\}$ при любом g; в самом деле, когда элемент g фиксирован, первое слагаемое во второй строке этих неравенств можно сделать сколь угодно малым, выбрав надлежащим образом элемент l, а второе будет стремиться к нулю при $m, n \to \infty$. Наконец, положив

$$\lim A_n g = Ag,$$

мы определим в пространстве L^2 функционал A, очевидно линейный, производящей функцией которого будет служить слабый предел f последовательности $\{f_{n_k}\}$. На этом доказательство теоремы выбора заканчивается.

33. Ортонормированные системы. Конечное или бесконечное множество $\{\varphi_n\}$ элементов пространства L^2 называется *ортогональной* и *нормированной* (или, короче, *ортонормированной*) *системой*, если

$$(\varphi_m, \varphi_n) = 0 \quad \text{при} \quad m \neq n \quad \text{и} \quad (\varphi_n, \varphi_n) = \|\varphi_n\|^2 = 1.$$

Рассмотрим сначала какую-нибудь конечную ортонормированную систему $\varphi_1, \varphi_2, \ldots, \varphi_N$. Возьмем произвольный элемент f пространства L^2 и попытаемся получить наилучшее приближение в среднем элемента f линейными комбинациями φ_k, т. е., варьируя коэффициенты c_k, свести к минимуму расстояние

$$\left\| f - \sum_{k=1}^{N} c_k \varphi_k \right\|.$$

В силу тождеств

$$\left\| f - \sum_{k=1}^{N} c_k \varphi_k \right\|^2 = \left(f - \sum_{k=1}^{N} c_k \varphi_k, \, f - \sum_{j=1}^{N} c_j \varphi_j \right) =$$

$$= (f, f) - \sum_{j=1}^{N} \bar{c}_j (f, \varphi_j) - \sum_{k=1}^{N} c_k (\varphi_k, f) + \sum_{k=1}^{N} c_k \bar{c}_k =$$

$$= \|f\|^2 - \sum_{k=1}^{N} |(f, \varphi_k)|^2 + \sum_{k=1}^{N} |(f, \varphi_k) - c_k|^2,$$

искомый минимум достигается при

$$c_k = (f, \varphi_k) \qquad (k = 1, 2, \ldots, N).$$

Введя эти значения в предыдущие соотношения, получим так называемое *тождество Бесселя*

$$\left\| f - \sum_{k=1}^{N} (f, \varphi_k) \varphi_k \right\|^2 = \|f\|^2 - \sum_{k=1}^{N} |(f, \varphi_k)|^2.$$

Так как его левая часть неотрицательна, то мы приходим к *неравенству Бесселя*

$$\sum_{k=1}^{N} |(f, \varphi_k)|^2 \leqslant \|f\|^2.$$

Из полученного неравенства

$$\left\| f-\sum_{k=1}^{N}(f,\ \varphi_k)\varphi_k \right\| \leqslant \left\| f-\sum_{k=1}^{N} c_k\varphi_k \right\|,$$

дающего решение рассмотрениой экстремальной задачи, следует, что если элемент f может быть с любой степенью точности приближен линейными комбинациями элементов $\varphi_1,\ \varphi_2,\ \ldots,\ \varphi_N$, то f сам является такой линейной комбинацией, а именно

$$f=\sum_{k=1}^{N}(f,\ \varphi_k)\varphi_k;$$

другими словами, множество всевозможных лииейных комбинаций $\sum_{k=1}^{N} c_k\varphi_k$ *замкнуто* относительно сходимости в среднем. Кроме того, это множество не совпадает со всем пространством L^2, иначе любые $N+1$ элементов пространства L^2 были бы линейно зависимы, что, очевидно, неверно.

Подмножество пространства L^2 мы назовем *полным* (в L^2), если оно порождает целиком все пространство, т. е. если произвольиый элемент L^2 может быть сколь угодно точно приближен в среднем элементами этого множества или их линейными комбинациями.

Мы показали, что конечная ортонормированная система не может быть полной в L^2. Для того чтобы установить существование в L^2 *счетной* полной ортонормированной системы, мы воспользуемся процессом ортогонализации, предложенным Шмидтом [1].

Возьмем в L^2 какую-нибудь последовательность (конечную или бесконечную, полную или неполную) элементов $f_n\neq 0$, каждый из которых линейно независим со всеми предшествующими. Сейчас мы *ортогонализируем* эту последовательность, т. е. заменим ее ортонормированной последовательностью $\{\varphi_n\}$, такой, что каждый элемент φ_n представляет собой линейную комбинацию элемеитов f_m с номерами $m\leqslant n$ и, обратно, каждый f_n есть линейная комбинация элементов φ_m с номерами $m\leqslant n$:

$$\begin{aligned} \varphi_n &= c_{n1}f_1+c_{n2}f_2+\ldots+c_{nn}f_n,\\ f_n &= \gamma_{n1}\varphi_1+\gamma_{n2}\varphi_2+\ldots+\gamma_{nn}\varphi_n. \end{aligned} \tag{10a}$$

Положим прежде всего $\varphi_1=\frac{1}{\|f_1\|}f_1$; тогда будут выполняться равенства $c_{11}^{-1}=\gamma_{11}=\|f_1\|$.

Теперь вычтем из f_2 элемент φ_1 с некоторым числовым множителем, подобранным так, чтобы разность

$$h_2=f_2-\gamma\varphi_1$$

была ортогональна φ_1:

$$(h_2,\ \varphi_1)=(f_2,\ \varphi_1)-\gamma(\varphi_1,\ \varphi_1)=0;$$

для этого надо взять $\gamma=(f_2, \varphi_1)$. Так как f_2 и f_1, а следовательно, также f_2 и φ_1 линейно независимы, то $h_2 \neq 0$, и мы положим

$$\varphi_2=\frac{1}{\|h_2\|}h_2.$$

При этом φ_2 оказывается линейной комбинацией элементов f_2 и φ_1, следовательно, линейной комбинацией f_2 и f_1; и, обратно, f_2 выражается линейно через h_2 и φ_1, т. е. через φ_2 и φ_1.

Вообще, если элементы $\varphi_1, \varphi_2, \ldots, \varphi_{r-1}$ определены так, что выполнены соотношения (10a) при $n=1, 2, \ldots, r-1$, то φ_r задается следующим образом. Числа λ_i выбираем так, чтобы элемент

$$h_r=f_r-\lambda_1\varphi_1-\lambda_2\varphi_2-\ldots-\lambda_{r-1}\varphi_{r-1}$$

был ортогонален элементам $\varphi_1, \varphi_2, \ldots, \varphi_{r-1}$, т. е. полагаем $\lambda_i=(f_r, \varphi_i)$, $i=1, 2, \ldots, r-1$. При этом $h_r \neq 0$, так как иначе f_r зависел бы линейно от $\varphi_1, \ldots, \varphi_{r-1}$ и, следовательно, от $f_1, \ldots, f_{r-1}$, что противоречит предположению. Можно, следовательно, положить

$$\varphi_r=\frac{1}{\|h_r\|}h_r.$$

Система $\varphi_1, \varphi_2, \ldots, \varphi_{r-1}, \varphi_r$ ортонормирована; φ_r представляет собой линейную комбинацию f_r и φ_i с номерами $i \leqslant r-1$, а следовательно, линейную комбинацию элементов f_i с $i \leqslant r$; обратно, f_r выражается линейно через h_r и φ_i с номерами $i \leqslant r-1$, а следовательно, через φ_i с $i \leqslant r$. Мы построили φ_r так, что выполняются соотношения (10a), и тем самым провели индукцию до конца.

Предположим теперь, что исходная последовательность $\{f_n\}$ *полная*. Полные последовательности существуют в силу того, что пространство L^2 сепарабельно: можно, например, взять последовательность характеристических функций всевозможных интервалов (R', R'') с рациональными концами, заключенных в основном интервале (a, b). Подвергнув $\{f_n\}$ ортогонализации, получим полную ортонормированную систему $\{\varphi_n\}$.

Систему, ортонормированную и полную в пространстве $L^2(0, 1)$, образуют, как известно из теории рядов Фурье, функции

$$e^{2\pi inx} \quad (n=0, \pm 1, \pm 2, \ldots).$$

Мы уже заметили, что никакая конечная ортонормированная система в L^2 не может быть полной. С другой стороны, L^2 не содержит несчетных ортонормированных систем. Это следует из того общего факта, что *любое подмножество сепарабельного метри-*

ческого пространства само сепарабельно[1]). Действительно, несчетная ортонормированная система не может быть сепарабельной, так как, в силу ортогональности, никакой ее элемент не может быть сколь угодно точно приближен линейными комбинациями остальных элементов.

Резюмируя, можно сказать, что L^2 представляет собой *счетномерное* пространство.

Рассмотрим теперь какую-нибудь счетную ортонормированную систему $\{\varphi_n\}$ и произвольный элемент f в пространстве L^2. Применив неравенство Бесселя к конечной системе $\varphi_1, \varphi_2, \ldots, \varphi_N$ и заставив N стремиться к бесконечности, получим в пределе

$$\sum_{k=1}^{\infty} |(f, \varphi_k)|^2 \leqslant \|f\|^2,$$

т. е. неравенство Бесселя для счетной ортонормированной системы. Ряд в левой части неравенства сходится, поэтому

$$\left\| \sum_{k=m}^{n} (f, \varphi_k)\varphi_k \right\|^2 = \sum_{k=m}^{n} |(f, \varphi_k)|^2 \to 0 \quad \text{при} \quad m, n \to \infty.$$

Отсюда, согласно теореме Рисса — Фишера, ряд

$$\sum_{k=1}^{\infty} (f, \varphi_k)\varphi_k \tag{106}$$

сходится в среднем к некоторому элементу g пространства L^2. Из сходимости в среднем вытекает слабая сходимость этого ряда, откуда

$$(g, \varphi_i) = \lim_{n \to \infty} \sum_{k=1}^{n} (f, \varphi_k)(\varphi_k, \varphi_i) = (f, \varphi_i);$$

---

[1]) Рассмотрим для определенности случай пространства L^2. Пусть $\{f_n\}$ — последовательность элементов, всюду плотная в L^2, а E — произвольное подмножество этого пространства. Для любой пары целых положительных чисел m, n выберем в E элемент g_{mn} (если таковой существует), для которого

$$\|g_{mn} - f_n\| < \frac{1}{m};$$

множество всех таких элементов, очевидно, счетно; покажем, что оно всюду плотно в E. Пусть g — произвольный элемент множества E и $\varepsilon > 0$. Выберем m так, чтобы выполнялось неравенство $\frac{1}{m} < \frac{\varepsilon}{2}$. При некотором n будет выполняться неравенство $\|g - f_n\| < \frac{1}{m}$, так как множество $\{f_n\}$ всюду плотно. Для указанных номеров m и n элемент g_{mn} существует и

$$\|g - g_{mn}\| \leqslant \|g - f_n\| + \|g_{mn} - f_n\| < \frac{1}{m} + \frac{1}{m} < \varepsilon,$$

что доказывает наше утверждение.

таким образом,

$$(f-g,\ \varphi_i)=0 \quad (i=1,\ 2,\ \ldots),$$

т. е. разность $f-g$ ортогональна всем φ_i, а следовательно, и всем линейным комбинациям этих элементов.

В том случае, когда рассматриваемая ортонормированная система *полна*, множество всевозможных линейных комбинаций ее элементов всюду плотно в L^2, и разность $f-g$ оказывается ортогональной любому элементу пространства L^2, в частности самой себе, т. е.

$$\|f-g\|^2=(f-g,\ f-g)=f-g=0,$$

откуда следует, что

$$f=g.$$

Нами доказана, таким образом,

Теорема. *Если $\{\varphi_n\}$ — полная ортонормированная система в L^2, то для любого элемента f пространства L^2 справедливо следующее разложение:*

$$f=\sum_{k=1}^{\infty}(f,\ \varphi_k)\,\varphi_k,$$

где ряд в правой части сходится в среднем.

Отсюда следует, что в случае полной ортонормированной системы левая часть тождества Бесселя стремится к нулю при $n\to\infty$ и вместо неравенства Бесселя мы получаем так называемую *формулу Парсеваля*

$$\sum_{1}^{\infty}|(f,\ \varphi_k)|^2=\|f\|^2.$$

То же название носит более общая формула

$$\sum_{1}^{\infty}(f_1,\ \varphi_k)\,\overline{(f_2,\ \varphi_k)}=(f_1,\ f_2),$$

в которой f_1 и f_2 означают любые два элемента пространства L^2. Вторая формула получается последовательным применением первой к элементам f_1+f_2, f_1-f_2, f_1+if_2, f_1-if_2 с последующим использованием следующего весьма полезного тождества:

$$4(f_1,\ f_2)=\|f_1+f_2\|^2-\|f_1-f_2\|^2+i\,\|f_1+if_2\|^2-i\,\|f_1-if_2\|^2.$$

Сходимость в среднем ряда (10б) была обусловлена лишь тем, что квадраты модулей коэффициентов этого ряда образовывали сходящийся ряд. Следовательно, имеет место следующая более общая

Теорема. *Каковы бы ни были ортонормированная (полная или нет) последовательность $\{\varphi_n\}$ в L^2 и числовая последовательность $\{a_n\}$, такая, что ряд*

$$\sum_{1}^{\infty}|a_n|^2$$

сходится, ряд

$$\sum_1^\infty a_n \varphi_n$$

сходится в среднем к некоторому элементу g пространства L^2, причем

$$(g, \varphi_n) = a_n \quad (n = 1, 2, \ldots).$$

Именно в такой форме теорема Рисса—Фишера была первоначально высказана Риссом.

34. Подпространства пространства L^2. Теорема о разложении. Подмножество E пространства L^2 называется *линейным множеством* (или *линейным многообразием*), если оно содержит всевозможные линейные комбинации своих элементов. Если E, кроме того, *замкнуто* относительно сходимости в среднем, то оно называется *подпространством* пространства L^2.

Конечномерное линейное множество, т. е. множество, образованное всевозможными линейными комбинациями

$$c_1 f_1 + c_2 f_2 + \ldots + c_N f_N$$

некоторых N фиксированных элементов $f_1, f_2, \ldots, f_N$, всегда *замкнуто, т. е. представляет собой подпространство пространства L^2*. Выше мы доказали это предложение для случая, когда $f_1, f_2, \ldots, f_N$ образуют ортонормированную систему; в общем случае его также можно доказать, изъяв из числа элементов f_i те, которые выражаются линейно через предшествующие, а затем ортогонализировав оставшуюся систему.

Следующая теорема весьма важна для приложений:

Теорема о разложении. *Пусть E—любое подпространство пространства L^2; каждый элемент h пространства L^2 может быть представлен, притом единственным образом, в виде суммы некоторого элемента f, принадлежащего E, и элемента g, ортогонального подпространству E (т. е. ортогонального любому элементу из E).*

Непосредственно очевидна *единственность* такого разложения; в самом деле, если $h = f + g$ и $h = f' + g'$ —два разложения, обладающие указанными свойствами, то элемент $f - f'$ принадлежит E и в то же время, будучи равен $g' - g$, ортогонален E, в частности ортогонален самому себе, откуда следует, что

$$\| f - f' \|^2 = (f - f', f - f') = 0$$

и $f = f'$, $g = g'$.

Существование такого разложения можно доказать, взяв ортонормированную последовательность элементов $\{\varphi_n\}$ в E, полную в E в том смысле, в каком раньше была определена полнота в L^2, и положив

$$f=\sum_n (h, \varphi_n)\varphi_n, \quad g=h-f;$$

систему $\{\varphi_n\}$ можно получить с помощью процесса ортогонализации из какой-либо последовательности, всюду плотной в E.

Приведем еще одно доказательство существования такого разложения, более прямое и *не использующее сепарабельность пространства* L^2 [1]).

Пусть h — фиксированный элемент из L^2, а f пробегает подпространство E. Рассмотрим „расстояния" $\|h-f\|$; пусть d — их нижняя грань, а $\{f_n\}$ — последовательность элементов из E, для которой

$$\|h-f_n\|\to d.$$

В равенстве (9) (см. п. 30) возьмем $h-f_m$ и $h-f_n$ соответственно вместо u и v; при этом вместо $\frac{1}{2}(u+v)$ будем иметь $h-\frac{f_m+f_n}{2}$. Так как одновременно с f_m и f_n их полусумма также принадлежит E, то

$$\left\|h-\frac{f_m+f_n}{2}\right\|\geqslant d.$$

Следовательно,

$$\|f_m-f_n\|^2\leqslant 2\|h-f_m\|^2+2\|h-f_n\|^2-4d^2\to 2d^2+2d^2-4d^2=0$$

и, согласно теореме Рисса — Фишера, последовательность $\{f_n\}$ сходится в среднем к некоторому элементу f^* подпространства E, для которого $\|h-f^*\|=d$. Пусть $g=h-f^*$; мы утверждаем, что g ортогонален любому элементу f из E. Для $f=0$ это очевидно; в общем случае воспользуемся неравенством

$$\|h-f^*\|^2\leqslant\|h-f^*-\lambda f\|^2=$$
$$=\|h-f^*\|^2-\bar\lambda(h-f^*, f)-\lambda(f, h-f^*)+\lambda\bar\lambda(f, f),$$

где положено

$$\lambda=\frac{(h-f^*, f)}{(f, f)}, \qquad \bar\lambda=\frac{(f, h-f^*)}{(f, f)}.$$

Мы получим

$$\frac{|(h-f^*, f)|^2}{\|f\|^2}\leqslant 0,$$

откуда следует, что $(h-f^*, f)=0$; это и требовалось доказать.

Доказательство существования экстремального элемента, только что изложенное, применимо к любому *выпуклому и замкнутому множеству* E, т. е. к такому множеству, которое, одновременно с f_1 и f_2, содержит все элементы вида $c_1f_1+c_2f_2$, где c_1, $c_2\geqslant 0$, $c_1+c_2=1$, а также содержит пределы последовательностей своих элементов, сходящихся в среднем.

---

[1]) См. Рисс [16]; по существу, здесь воспроизводится построение, осуществленное ранее Б. Леви (см. [2], § 7) в его исследованиях, касающихся принципа Дирихле.

Отметим одно из следствий теоремы:

Если в пространстве L^2 или в каком-либо его подпространстве F дано множество E элементов f, линейные комбинации которых не всюду плотны в F в смысле сходимости в среднем, то в F найдутся элементы $g \neq 0$, ортогональные множеству E.

Теорему о разложении можно сформулировать еще и так:

Пусть в L^2 задано множество E элементов, линейные комбинации которых не всюду плотны в L^2 в смысле сходимости в среднем. Элементы, ортогональные множеству E, образуют некоторое подпространство, а элементы, ортогональные этому последнему,— другое подпространство, которое состоит из линейных комбинаций элементов, принадлежащих E, и их пределов и представляет собой наименьшее подпространство, содержащее E; оно называется подпространством, порожденным множеством E или натянутым на множество E. Любой элемент f из L^2 представляется, притом единственным образом, в виде $f = h + g$, где h и g принадлежат соответственно обоим взаимноортогональным „дополнительным“ подпространствам.

35. Другое доказательство теоремы выбора. Продолжение функционалов. Отметим два интересных следствия теоремы о разложении. Первое — доказательство теоремы выбора, отличающееся от приведенного в п. 32 тем, что оно не опирается на сепарабельность пространства L^2.

Пусть $\{f_n\}$ — ограниченная последовательность элементов из L^2. С помощью диагонального метода выделим подпоследовательность $\{f_{n_k}\}$, такую, что скалярные произведения (f_m, f_{n_k}) сходятся при любом фиксированном m. Легко видеть, что и при произвольном f скалярные произведения (f, f_{n_k}) будут сходиться. Это очевидно в том случае, когда f есть линейная комбинация элементов f_m или предел таких линейных комбинаций. Если элементы f такого вида не исчерпывают всего L^2, то, представив f в виде $f = g + h$, где g принадлежит подпространству, порожденному последовательностью $\{f_n\}$, а h ортогонально этому подпространству и, в частности, всем f_{n_k}, мы получим

$$(f, f_{n_k}) = (g, f_{n_k}) + (h, f_{n_k}) = (g, f_{n_k}),$$

откуда будет следовать сходимость скалярных произведений (f, f_{n_k}).

Второе следствие теоремы о разложении, которое мы имеем в виду, относится к проблеме продолжения линейных функционалов.

Теорема. *Пусть A — функционал, заданный лишь на некотором множестве E; предположим, что существует число M,* *такое, что для всех линейных комбинаций* $\sum_1^n c_k f_k$ *элементов* f_k *из* E *выполняется неравенство*

$$\left|\sum_{k=1}^{n} c_k A f_k\right| \leqslant M \left\|\sum_{k=1}^{n} c_k f_k\right\|.$$

Тогда функционал A *может быть продолжен на все пространство* L^2 *как линейный функционал с нормой* M_A, *не превосходящей* M.

Требуемый линейный функционал определяется сначала на линейных комбинациях элементов из E посредством равенства

$$A\left(\sum_1^n c_k f_k\right) = \sum_1^n c_k A f_k,$$

затем на их пределах очевидным предельным переходом и, наконец, полагается равным нулю на элементах, ортогональных множеству E. Пользуясь разложением $f = g + h$, мы полагаем $Af = Ag + Ah$ и тем самым определяем Af для произвольного f.

36. Пространство L^p и его линейные функционалы. Рассмотрим теперь пространства L^p $(p \geqslant 1)$, образованные измеримыми функциями f, заданными на некотором измеримом множестве e, для которых $|f|^p$ суммируемы. Норма $\|f\|$ в L^p определяется равенством

$$\|f\| = \left(\int_e |f(x)|^p dx\right)^{\frac{1}{p}};$$

при этом, как и в случае L^2, $\|f\| = 0$ тогда и только тогда, когда $f(x) = 0$ почти всюду. Так же, как в L^2, две функции, совпадающие почти всюду, удобно считать тождественными.

Неравенство Минковского, иначе говоря, неравенство треугольника в пространстве L^p, справедливо, как мы видели, при любом $p > 1$; очевидно, оно распространяется и на $p = 1$, а также на предельный случай $p = \infty$. Что касается этого последнего, то под L^∞ понимается совокупность измеримых функций f, каждая из которых ограничена или совпадает почти всюду с ограниченной функцией, причем норма $\|f\|$ определяется как „существенная верхняя грань" $|f(x)|$, т. е. наименьшее из чисел M, для которых неравенство $|f(x)| \leqslant M$ выполняется почти всюду. Точка зрения на это пространство как на предельный случай пространств L^p оправдывается тем, что, как легко видеть, в случае множества e конечной меры нормы функции f в L^p как функции от p стремятся при $p \to \infty$ к норме в пространстве L^∞.

Чтобы определить скалярное произведение, надо, помимо L^p, рассмотреть пространство L^q, где $\frac{1}{p} + \frac{1}{q} = 1$, т. е. $q = \frac{p}{p-1}$, $p =$

$=\frac{q}{q-1}$; предельным случаям $p=1$ и $p=\infty$ отвечают значения $q=\infty$ и $q=1$. Скалярное произведение определяется так же, как в L^2, но в данном случае один из множителей должен принадлежать L^p, а другой L^q.

Для упрощения записи мы будем рассматривать действительные пространства L^p, образованные действительными функциями; переход к комплексному случаю читатель осуществит без труда.

Сходимость в среднем при $p \geqslant 1$ определяется в точности так же, как при $p=2$, только под $\|f-f_n\|$ и $\|f_m-f_n\|$ следует, разумеется, понимать нормы в L^p. Критерий сходимости, соответствующий теореме Рисса—Фишера, короче, *критерий Коши* устанавливается точно так же, как в L^2, а в частном случае $p=1$ даже проще. В предельном случае $p=\infty$ действует классический критерий, так как мы имеем дело, по существу, с равномерной сходимостью. Следует заметить, что наши определения применимы и в случае положительных $p<1$, но обобщение в этом направлении оказывается мало интересным[1]).

Что касается понятия *слабой сходимости* в пространстве L^p, $1 \leqslant p < \infty$, то последовательность $\{f_n\}$ в L^p мы назовем слабо сходящейся к f, если для любого элемента g „сопряженного" пространства L^q

$$(f_n, g)=\int_e f_n(x)\, g(x)\, dx \to (f, g).$$

Так же, как в случае $p=2$, слабую сходимость можно определить с помощью линейных функционалов. Действительно, скалярное произведение (f, g) при фиксированном f из L^p и переменном g из L^q определяет в L^q линейный функционал, обладающий свойствами, аналогичными перечисленным в п. 30. Это очевидно; менее очевидно обратное предложение, состоящее в том, что при $1 \leqslant q < \infty$ всякий линейный функционал в L^q (в действительных пространствах мы рассматриваем лишь действительные функционалы) может быть представлен в виде (f, g), где производящая функция $f(x)$ принадлежит пространству L^p.

Прежде чем доказывать это, установим необходимое и достаточное условие того, что заданная функция $F(x)$, определенная в конечном или бесконечном интервале (a, b), является неопределенным интегралом некоторой функции $f(x)$, принадлежащей L^p.

Лемма[2]). *Функция $F(x)$ представляет собой неопределенный интеграл некоторого элемента $f(x)$ пространства L^p $(1<p<\infty)$ тогда и только тогда, когда при любом выборе конечной системы*

---

[1]) При $p<1$ в L^p не существует линейных функционалов, не равных нулю тождественно (см. Дэй [1]).

[2]) Рисс [6] (§ 5).

точек

$$x_0 < x_1 < \ldots < x_{m-1} < x_m$$

в интервале (a, b) *суммы*

$$\sum_{1}^{m} \frac{|F(x_k)-F(x_{k-1})|^p}{(x_k-x_{k-1})^{p-1}} \tag{11}$$

равномерно ограничены сверху. Верхняя грань таких сумм совпадает с интегралом функции $|f(x)|^p$ *на* (a, b).

Сначала заметим, что эта лемма, коль скоро она верна для любого конечного интервала, непосредственно распространяется на случай бесконечных интервалов. Поэтому мы предположим, что интервал (a, b) конечен.

Необходимость условия вытекает из неравенства Гёльдера, примéненного к функциям $f(x)$ и 1:

$$|F(x_k)-F(x_{k-1})|^p \leqslant \left(\int_{x_{k-1}}^{x_k} |f(x)|\,dx\right)^p \leqslant$$

$$\leqslant (x_k-x_{k-1})^{p-1} \int_{x_{k-1}}^{x_k} |f(x)|^p\,dx. \tag{11a}$$

Допустим теперь, что суммы (11) ограничены; пусть B^p — их верхняя грань. Сумма, подобная (11), но построенная для системы непересекающихся интервалов (α_k, β_k) (но не покрывающих, вообще говоря, интервал (a, b)), тем более не превзойдет B^p; отсюда, воспользовавшись неравенством Гёльдера для конечных сумм [1]), получим

$$\sum |F(\beta_k)-F(\alpha_k)| \leqslant$$

$$\leqslant \left(\sum \frac{|F(\beta_k)-F(\alpha_k)|^p}{(\beta_k-\alpha_k)^{p-1}}\right)^{\frac{1}{p}} \left(\sum (\beta_k-\alpha_k)\right)^{\frac{p-1}{p}} \leqslant B \left(\sum (\beta_k-\alpha_k)\right)^{\frac{p-1}{p}}.$$

Таким образом, функция $F(x)$ абсолютно непрерывна и, следовательно, имеет почти всюду производную $F'(x)$, неопределенным интегралом которой она является. Эта производная служит пределом последовательности ступенчатых функций $f_n(x)$, которые можно построить с помощью последовательности разбиений интервала (a, b), например на 2^n равных частей (α, β), положив, что на каждой такой части $f_n(x)$ равна

$$\frac{F(\beta)-F(\alpha)}{\beta-\alpha}.$$

---

[1]) $\left|\sum a_k b_k\right| \leqslant \left(\sum |a_k|^p\right)^{\frac{1}{p}} \left(\sum |b_k|^q\right)^{\frac{1}{q}} \left(p > 1,\ \frac{1}{p}+\frac{1}{q}=1\right).$

Ясно, что сумма (11), соответствующая такому разбиению, будет равна интегралу функции $|f_n|^p$ на (a, b). Согласно лемме Фату, функция $|F'(x)|^p$ суммируема и ее интеграл не превосходит B^p. С другой стороны, в силу (11а), суммы (11), а следовательно, и их верхняя грань B^p не превосходят интеграла $|F'(x)|^p$ на (a, b); итак, этот интеграл равен B^p, что и требовалось доказать.

Вернемся к теореме о представлении линейного функционала в виде интеграла. Не нарушая общности, мы можем ограничиться случаем, когда измеримое множество e, на котором определены рассматриваемые функции, представляет собой конечный или бесконечный интервал (a, b)[1]. Пусть Ag — линейный функционал в пространстве L^q $(1 \leqslant q < \infty)$; положим $F(x)$ равной Ag_x, где $g_x(\xi)$ — характеристическая функция интервала (a, x) (в случае $a = -\infty$, чтобы избежать осложнений, связанных с неинтегрируемостью функции, $g_x(\xi)$ определяется как характеристическая функция интервала $(0, x)$ при $x > 0$, а при $x < 0$ — как характеристическая функция интервала $(x, 0)$, умноженная на -1).

Рассмотрим сначала случай $1 < q < \infty$. Покажем, что функция $F(x)$ удовлетворяет условию леммы при значении $p = \frac{q}{q-1}$ и, следовательно, является неопределенным интегралом некоторой функции $f(x)$, принадлежащей L^p. Для этого возьмем ступенчатую функцию $\varphi(\xi)$, принимающую на интервалах (x_{k-1}, x_k) постоянные значения[2]:

$$\frac{|F(x_k) - F(x_{k-1})|^{p-1} \operatorname{sgn}(F(x_k) - F(x_{k-1}))}{(x_k - x_{k-1})^{p-1}}.$$

При этом

$$A\varphi = \sum_1^m \frac{|F(x_k) - F(x_{k-1})|^p}{(x_k - x_{k-1})^{p-1}};$$

---

[1]) Действительно, если e — измеримое подмножество интервала (a, b), то любой линейный функционал в пространстве L^q функций, определенных на e, может рассматриваться как линейный функционал в пространстве L^q функций, определенных на интервале (a, b), который определяется лишь значениями этих функций, принимаемыми на e. Для производящей функции f этого функционала и любой функции $g(x)$ из L^q, тождественно равной нулю на e, будет выполняться соотношение

$$\int_a^b g(x) f(x)\, dx = 0.$$

Отсюда вытекает, что функция $f(x)$ равна нулю всюду вне e.

[2]) Функция $\operatorname{sgn} a$ принимает значения 1, 0 и -1 соответственно при $a > 0$, $a = 0$ и $a < 0$.

с другой стороны, так как $q(p-1)=q\left(\frac{q}{q-1}-1\right)=p$, то

$$A\varphi \leqslant M_A \|\varphi\| = M_A\left(\int_a^b |\varphi(\xi)|^q d\xi\right)^{\frac{1}{q}} =$$

$$= M_A\left(\sum_1^m \frac{|F(x_k)-F(x_{k-1})|^p}{(x_k-x_{k-1})^{p-1}}\right)^{\frac{1}{q}},$$

откуда следует, что

$$\sum_1^m \frac{|F(x_k)-F(x_{k-1})|^p}{(x_k-x_{k-1})^{p-1}} \leqslant M_A^p.$$

Итак, $F(x)$ представляет собой неопределенный интеграл некоторой функции $f(x)$, принадлежащей пространству L^p, причем

$$\int_a^b |f(x)|^p dx \leqslant M_A^p.$$

Соотношение

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

очевидно для ступенчатых функций g, а так как эти последние всюду плотны в L^q, то оно выполняется для произвольного g из L^q. Неравенство Гёльдера дает нам

$$|Ag|=\left|\int_a^b g(x) f(x)\, dx\right| \leqslant \left(\int_a^b |g(x)|^q dx\right)^{\frac{1}{q}} \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}},$$

откуда следует, что

$$M_A^p \leqslant \int_a^b |f(x)|^p dx;$$

сопоставляя последнее неравенство с полученным выше, мы придем к выводу, что M_A^p равно этому интегралу.

Рассмотрим отдельно более простой случай $p=1$. Для любой пары точек x_1, x_2 интервала (a, b)

$$|F(x_1)-F(x_2)| = |A(g_{x_1}-g_{x_2})| \leqslant$$

$$\leqslant M_A \int_a^b |g_{x_1}(\xi)-g_{x_2}(\xi)|\, d\xi = M_A |x_2-x_1|,$$

т. е. $F(x)$ удовлетворяет условию Липшица с постоянной M_A. Отсюда следует (см. п. 25), что $F(x)$ является неопределенным интегралом некоторой функции $f(x)$, причем $|f(x)| \leqslant M_A$. Так же, как в предыдущем случае, доказывается, что

$$Ag = \int_a^b g(x) f(x)\, dx = (g, f).$$

Отсюда вытекает, что

$$|Ag| \leqslant (\text{vrai}\max |f(x)|) \int_a^b |g(x)|\, dx$$

и, следовательно,

$$M_A \leqslant \text{vrai}\max |f(x)|;$$

а так как $|f(x)| \leqslant M_A$, то $M_A = \text{vrai}\max |f(x)|$.

Поменяв местами p и q, а также f и g, мы получим следующий результат:

Теорема[1]). *Всякому линейному функционалу Af в пространстве L^p $(1 \leqslant p < \infty)$ соответствует производящая функция $g(x)$, определенная почти всюду и принадлежащая пространству L^q, посредством которой функционал A представляется в виде*

$$Af = (f, g);$$

при этом норма функционала равна норме производящей функции в соответствующем пространстве:

$$M_A = \begin{cases} \left(\int\limits_e |g(x)|^q dx\right)^{\frac{1}{q}} & \text{при } p > 1,\ q < \infty, \\ \text{vrai}\max |g(x)| & \text{при } p = 1,\ q = \infty. \end{cases}$$

Мы покажем ниже (см. п. 87), что для $p = \infty$ эта теорема неверна: в пространстве L^∞ существуют линейные функционалы, которые не порождаются функциями, принадлежащими L^1.

Только что доказанная теорема позволяет почти дословно перенести на пространства L^p как формулировку, так и доказательство теоремы выбора (см. п. 32).

37. Одна теорема о сходимости в среднем. Приведем еще одну теорему, которая, как мы видели выше (см. п. 29), при $p = 2$ доказывается с помощью совсем простой выкладки, в общем же случае потребует более тонких рассмотрений.

Теорема[2]). *Если последовательность $\{f_n\}$ в пространстве L^p $(1 < p < \infty)$ сходится слабо к некоторому элементу f и если, кроме того,*

$$\|f_n\| \to \|f\|,$$

то $\{f_n\}$ сходится к f в среднем.

---

[1]) Рисс [6], § 11, [23].

[2]) Радон [1] (стр. 1363); см. также Рисс [13].

Предположим сначала, что $p \geqslant 2$; тогда, каково бы ни было действительное число z, будет выполняться неравенство

$$|1+z|^p \geqslant 1+pz+c|z|^p, \tag{12}$$

где c — некоторая положительная постоянная, не зависящая от z. В самом деле, рассмотрим дробь

$$\frac{|1+z|^p-1-pz}{|z|^p}; \tag{13}$$

вторая производная числителя

$$p(p-1)|1+z|^{p-2}$$

положительна и только при $z=-1$ обращается в нуль, а так как сам числитель и его первая производная обращаются в нуль при $z=0$, то $|1+z|^p-1-pz>0$, когда $z \neq 0$. При $z=0$ дробь (13) обращается в бесконечность, если $p>2$, и становится равной 1, если $p=2$; следовательно, она положительна при всех значениях z; наконец, при $|z| \to \infty$ дробь (13) стремится к 1, следовательно, она достигает в какой-то точке своего наименьшего значения c, причем $c>0$.

В неравенстве (12) положим

$$z=\frac{f_n(x)-f(x)}{f(x)};$$

умножив на $|f|^p$ и проинтегрировав, получим

$$\int_e |f_n|^p dx \geqslant \int_e |f|^p dx + p\int_e |f|^{p-2} f(f_n-f)\,dx + c\int_e |f_n-f|^p dx.$$

Так как $\{f_n\}$ сходится слабо к f, то второй интеграл в правой части стремится к нулю одновременно с $\frac{1}{n}$, а, в силу предположения относительно норм, левая часть стремится к первому интегралу справа; следовательно,

$$\int_e |f_n-f|^p dx \to 0, \tag{14}$$

что и требовалось доказать.

В случае, когда $1<p<2$, рассмотрим вместо (13) функцию, равную (13) при $|z| \geqslant 1$, а при $|z| \leqslant 1$ равную

$$\frac{|1+z|^p-1-pz}{z^2}.$$

В точке $z=0$ эта функция имеет предел $\frac{1}{2}p(p-1)$, и так как она положительна при всех $z \neq 0$, конечных и бесконечных, то она тоже имеет положительное наименьшее значение. Отсюда, рассуждая так же, как в первом случае, мы получим вместо требуемого соотношения (14) лишь соотношение

$$\int_{e_n} |f_n - f|^p\,dx + \int_{e \setminus e_n} (f_n - f)^2 |f|^{p-2}\,dx \to 0,$$

где e_n означает множество тех значений x, для которых

$$|f_n(x) - f(x)| \geqslant |f(x)|.$$

Оба последних интеграла стремятся к нулю; соотношение (14) будет доказано, если мы установим, что из

$$\int_{e \setminus e_n} (f_n - f)^2 |f|^{p-2}\,dx \to 0$$

следует

$$\int_{e \setminus e_n} |f_n - f|^p\,dx \to 0.$$

Чтобы это обнаружить, воспользуемся неравенством Шварца, а также неравенством $|f_n(x) - f(x)| < |f(x)|$, справедливым на множестве $e \setminus e_n$:

$$\int_{e \setminus e_n} |f_n - f|^p\,dx \leqslant \int_{e \setminus e_n} |f|^{p-1} |f_n - f|\,dx \leqslant$$

$$\leqslant \left(\int_{e \setminus e_n} |f|^p\,dx\right)^{\frac{1}{2}} \left(\int_{e \setminus e_n} (f_n - f)^2 |f|^{p-2}\,dx\right)^{\frac{1}{2}} \leqslant$$

$$\leqslant \left(\int_{e} |f|^p\,dx\right)^{\frac{1}{2}} \left(\int_{e \setminus e_n} (f_n - f)^2 |f|^{p-2}\,dx\right)^{\frac{1}{2}} \to 0.$$

Итак, теорема доказана и для $p < 2$.

38. Теорема Банаха — Сакса. К рассматриваемому кругу идей относится следующая

Теорема. *Если последовательность $\{f_n\}$ в L^p слабо сходится к f, то из нее можно выбрать подпоследовательность $\{f_{n_k}\}$, такую, что средние арифметические*

$$\frac{1}{k}(f_{n_1} + f_{n_2} + \ldots + f_{n_k})$$

сходятся к f в среднем.

Эта теорема доказана польскими математиками Банахом и Саксом[1]), чьи труды в вопросах, которым посвящена эта книга,

[1]) Банах и Сакс [1]. Эта теорема доказана также для функционального пространства C, которое мы рассмотрим в гл. III, и даже для произвольного „банахова пространства" (см. Мазур [1]).

получили всеобщее признание. Предложенное этими авторами доказательство близко к тому, которое изложено здесь, но несколько более громоздко в подробностях; отсылая читателя к статье, упомянутой в примечании, мы ограничимся лишь случаем $p=2$.

Мы можем предположить, что $f=0$, так как в противном случае мы взяли бы f_n-f вместо f_n. Положим для определенности $n_1=1$; пусть n_2—такой номер (или, если угодно, наименьший из таких номеров), для которого $|(f_1, f_{n_2})|\leqslant 1$; такие номера существуют, так как $(f_1, f_n)\to 0$. Если $f_{n_1}, f_{n_2}, \ldots, f_{n_k}$ уже выбраны, то номер $n_{k+1}>n_k$ мы выберем так, чтобы выполнялись неравенства

$$|(f_{n_1}, f_{n_{k+1}})|\leqslant\frac{1}{k}, \ \ldots, \ |(f_{n_k}, f_{n_{k+1}})|\leqslant\frac{1}{k};$$

такой выбор возможен в силу того, что $(f_{n_i}, f_n)\to 0$ $(i=1, 2, \ldots, k;$ $n\to\infty)$. Далее, нормы $\|f_n\|$ ограничены,

$$\|f_n\|\leqslant B,$$

следовательно,

$$\left\|\frac{1}{k}(f_{n_1}+f_{n_2}+\ldots+f_{n_k})\right\|^2\leqslant\frac{kB^2+2\cdot 1+4\cdot\frac{1}{2}+\ldots+2(k-1)\frac{1}{k-1}}{k^2}<$$
$$<\frac{B^2+2}{k}\to 0,$$

что и требовалось доказать. Коротко можно сказать, что каждый элемент f_{n_k} выбирается таким образом, чтобы он был приблизительно ортогонален ранее выбранным f_{n_i}.

Из доказанной теоремы вытекает следующий важный результат.

Теорема. *Если линейное (или хотя бы выпуклое) множество в L^p замкнуто в смысле сходимости в среднем, то оно замкнуто и в смысле слабой сходимости.*

§ 4. ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

39. Определения. Принцип соответствия. Мы ограничимся случаем двух переменных; этого достаточно для выяснения общей картины.

Для построения теории интегрирования нам понадобится шаг за шагом воспроизводить определения и построения, которыми мы пользовались при рассмотрении функций одного переменного; исходным материалом должны служить ступенчатые функции двух переменных, т. е. функции, принимающие постоянные значения на некоторых прямоугольниках (или даже квадратах) со сторонами, параллельными координатным осям. Оперируя с такими прямоугольниками вместо интервалов, мы сможем значи-