ЦИКЛ СТАТЕЙ ПО ФУНКЦИОНАЛЬНОМУ АНАЛИЗУ.

ОСНОВНЫЕ ПОНЯТИЯ ФУНКЦИОНАЛЬНОГО АНАЛИЗА.

Л. А. Люстерник.

§ 1. Введение.

Настоящая статья представляет собой обработку лекций, читанных автором в Воронежском университете весной 1935 г. и посвященных наиболее общим понятиям современного анализа и геометрии.

Основное понятие анализа есть понятие *функциональной зависимости*:

$$y = f(x). \tag{1}$$

Мы относим, в простейшем случае, каждому вещественному числу (аргументу) x такое же число (функцию) y. В функциях многих переменных аргументом является группа вещественных чисел, или, допуская геометрическую интерпретацию, — точка n-мерного пространства. В вариационном исчислении мы уже имеем значительно более общий вид функциональной зависимости. В качестве аргументов здесь фигурируют линии, поверхности и т. п. Например, длина кривой есть величина, зависящая от кривой. Мы можем в формуле (1) под x понимать кривую, а под y — вещественное число, ее длину. Мы приходим к понятию *функционала*: дан класс X некоторых элементов. Каждому элементу x из X относится вещественное число y — функционал от x.

Можно итти дальше в обобщении понятия функциональной зависимости, рассматривая и y в левой части (1) как элементы весьма общей природы. В вектор-функциях мы в качестве y имеем n-мерные векторы. Встречающийся в теории интегральных уравнений оператор Фредгольма

$$y(s) = \int K(s, t)\, x(t)\, dt \tag{2}$$

относит каждой непрерывной функции $x(t)$ непрерывную же функцию $y(t)$; выражение (2) можно рассматривать как частный случай соотношения (1), когда в качестве y и x фигурируют непрерывные функции.

Мы приходим к следующему, наиболее общему представлению о функциональной зависимости: даны два множества элементов X и Y. Каждому элементу x из X мы относим по некоторому закону элемент y из Y. Выражение (1) есть символ этого соответствия. Но, очевидно, относительно таких общих соответствий можно сказать очень немного. Для того чтобы эти соответствия сохраняли в некоторой степени те свойства функций, которые рассматриваются в обычном анализе, нужно, чтобы на множества X и Y распространялись некоторые свойства объектов обычного анализа.

При построении анализа основным понятием является понятие *предельного перехода*. Естественно ограничиться теми множествами X и Y, для элементов которых установлены соотношения, аналогичные предельному переходу для вещественных чисел. Заметим, что в математике наряду с предельным переходом для числовых последовательностей изучаются другие предельные переходы, например для последовательностей функций (сходимость равномерная, сходимость в среднем, сходимость по мере, сходимость в каждой точке, сходимость почти всюду и т. п.).

Совокупности элементов, для которых установлены законы предельного перехода, обобщающие основные законы предельного перехода для числовых последовательностей, называются *абстрактными пространствами*.

Существует целый ряд аксиоматик абстрактных пространств. Мы здесь остановимся на одном классе абстрактных пространств, так называемых метрических пространствах [1]). *Метрическим пространством* называется множество M, между каждой парой элементов которого установлено некоторое числовое соотношение — *расстояние*. Мы требуем, чтобы расстояние между элементами a и b из M, которое мы будем обозначать через $\rho(a, b)$, удовлетворяло некоторым аксиомам:

1. $\rho(a, b)$ есть неотрицательное число, $\rho(a, b) \geqslant 0$, причем $\rho(a, b) = 0$ в том и только в том случае, если a совпадает с b, $a = b$.

2. $\rho(a, b) = \rho(b, a)$.

3. $\rho(a, b) + \rho(b, c) \geqslant \rho(a, c)$ (аксиома треугольника).

Аксиомы метрики 1—3 описывают основные свойства расстояний между точками обычного пространства.

Мы оправдываем геометрическую терминологию: при создании абстрактных пространств мы сохраняем некоторые из наиболее общих свойств обычного пространства, в данном случае — свойства расстояний.

Элементы метрического пространства будем называть его *точками*. Предельный переход определяется в метрических пространствах совершенно естественно. Будем говорить: a_n стремится к a и писать

$$a_n \to a,$$

если

$$\rho(a_n, a) \to 0.$$

Примеры. 1. Простейшим примером может служить евклидово n-мерное пространство E_n, точка которого $x(x_1, x_2, \ldots, x_n)$ есть совокупность n чисел $x_1, x_2, \ldots, x_n$ — „координат" x, и расстояние между точками $x(x_1, x_2, \ldots, x_n)$ и $y(y_1, y_2, \ldots, y_n)$ которого определяется по формуле

$$\rho(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}. \tag{3}$$

2. Можно определить расстояние в n-мерном пространстве иначе, например

$$\rho(x, y) = \left\{\sum_{i=1}^{n} |x_i - y_i|^p\right\}^{\frac{1}{p}} \tag{4}$$

---

[1]) По поводу теории метрических, как и вообще абстрактных пространств, см. книгу Хаусдорфа, Теория множеств (готовится к печати).

при $p \geqslant 1$. Определенное таким образом расстояние удовлетворяет аксиомам метрики (что вытекает из неравенства Минковского) [1]); такое пространство будем называть $l_p^{(n)}$.

3. Гильбертово пространство определяется как совокупность числовых последовательностей $x(x_1, x_2, \ldots)$, для которых $\sum_{i=1}^{\infty} x_i^2$ сходится, причем метрика определяется по аналогии с евклидовым пространством: если $x(x_1, x_2, \ldots)$ и $y(y_1, y_2, \ldots)$ суть точки гильбертова пространства, то

$$\rho(x, y) = \sqrt{\sum_{i=1}^{\infty} (x_i - y_i)^2}. \tag{5}$$

4. Пространство l_p ($p \geqslant 1$) есть совокупность всех числовых последовательностей $(x_1, x_2, \ldots)$, для которых $\sum_{i=1}^{\infty} |x_i|^p$ сходится. Расстояние между $x(x_1, x_2, \ldots)$ и $y(y_1, y_2, \ldots)$ в l_p определяется по формуле

$$\rho(x, y) = \left\{ \sum_{i=1}^{\infty} |x_i - y_i|^p \right\}^{\frac{1}{p}} \tag{6}$$

(l_2 есть гильбертово пространство).

l_p дает пример координатного пространства: его элемент $x(x_1, x_2, \ldots)$ определяется счетным множеством чисел — координат x_i. Мы будем писать вместо $(x_1, x_2, \ldots)$ коротко: (x_i).

5. Пространство m есть множество всех ограниченных последовательностей вещественных чисел, причем для $x(x_i)$ и $y(y_i)$ из m расстояние определяется формулой

$$\rho(x, y) = \sup |x_i - y_i|.$$

6. N есть совокупность всех числовых последовательностей $x(x_i)$. Определим сходимость в N следующим образом: $x^{(k)}(x_i^{(k)})$, $k = 1, 2, 3, \ldots,$ сходится к $x(x_i)$, если одновременно

$$x_i^{(k)} \to x_i \quad (k \to \infty), \quad i = 1, 2, 3, \ldots$$

N есть абстрактное неметрическое пространство.

---

[1]) Неравенством Минковского называется неравенство

$$\left\{ \sum_{i=1}^{n} |x_i + y_i|^p \right\}^{\frac{1}{p}} \leqslant \left\{ \sum_{i=1}^{n} |x_i|^p \right\}^{\frac{1}{p}} + \left\{ \sum_{i=1}^{n} |y_i|^p \right\}^{\frac{1}{p}} \quad (p \geqslant 1).$$

Это неравенство переходит в равенство, если все x_i пропорциональны y_i. Аналогичное неравенство имеет место для интегралов:

$$\left\{ \int |y + z|^p \, dt \right\}^{\frac{1}{p}} \leqslant \left\{ \int |y|^p \, dt \right\}^{\frac{1}{p}} + \left\{ \int |z|^p \, dt \right\}^{\frac{1}{p}} \quad (p \geqslant 1).$$

См., например, Лаврентьев и Люстерник, Элементы вариационного исчисления, т. I, ч. 1, стр. 83—84.

7. Пространство L_p, $p \geqslant 1$, есть совокупность всех измеримых функций $y(t)$, $0 \leqslant t \leqslant 1$, для которых существует интеграл

$$\int_0^1 |y(t)|^p dt.$$

Метрика в L_p определяется формулой

$$\rho(y, z) = \left\{ \int_0^1 |y - z|^p dt \right\}^{\frac{1}{p}}. \tag{7}$$

Сходимость y_n к y в L_p означает стремление к нулю выражения

$$\int_0^1 |y_n(t) - y(t)|^p dt$$

(т. е. сходимость в среднем p-й степени). В частности, при $p = 2$, мы получаем обычную сходимость в среднем. Заметим, что две функции $y(t)$, $\bar{y}(t)$, различающиеся на множестве меры 0, должны считаться совпадающими в L_p (для них $\rho(y, \bar{y}) = 0$).

8. Пространство C состоит из всех непрерывных функций $y(t)$, $0 \leqslant t \leqslant 1$. Расстояние в C определяется формулой

$$\rho(y, z) = \max_{0 \leqslant t \leqslant 1} |y(t) - z(t)|. \tag{8}$$

Расстояние в C совпадает с так называемой близостью нулевого порядка в вариационном исчислении. Сходимость в C, как легко убедиться, есть равномерная сходимость.

9. Пространство Q. Рассмотрим совокупность плоских кривых q, заданных уравнениями

$$\left.\begin{array}{l} x = x(t) \\ y = y(t) \end{array}\right\} \quad (0 \leqslant t \leqslant 1). \tag{9}$$

Если заданы две кривые q и q_1, то, записав их уравнения в параметрическом виде (9), будем относить друг другу точки обеих кривых, отвечающие одинаковым значениям параметра t. Обозначим через ε максимум расстояний между соответственными точками обеих кривых. Число ε зависит от выбора параметрического представления кривых. Будет считать расстояние

$$\rho(q, q_1)$$

равным нижней грани чисел ε при всевозможных параметрических представлениях обеих кривых.

Введенная в Q метрика, как легко убедиться, удовлетворяет всем аксиомам метрики.

10. Пространство S состоит из всех измеримых на отрезке $[0, 1]$ функций $y(t)$; для каждой $y(t) \subset S$ определим множестве E_y чисел ε, для каждого из которых неравенство

$$|y(t)| > \varepsilon$$

удовлетворяется на множестве M_ε, причем $\operatorname{mes} M_\varepsilon \leqslant \varepsilon$. Например, число 1 входит в E_y, ибо множество точек $t \subset [0, 1]$, для которых $|y(t)| > 1$, имеет меру во всяком случае не бóльшую единицы (длины всего отрезка $[0, 1]$).

Обозначим через $\varepsilon_0(y)$ нижнюю грань чисел $\varepsilon \subset E_y$. Определим метрику в S следующим образом: $\rho[x(t), y(t)] = \varepsilon_0(x-y)$. Для любого $\varepsilon > \rho(x, y)$ неравенство $|x(t)-y(t)| > \varepsilon$ имеет место на множестве, по мере не превосходящем ε. Сходимость в S есть сходимость по мере.

Операторы и функционалы. Пусть даны два абстрактных пространства M и M_1. Пусть дано соответствие (1):

$$y = f(x),$$

относящее каждому $x \subset M$ элемент $y \subset M_1$. Такое соотношение будем называть *оператором*, отображающим M на M_1.

Оператор называется *непрерывным*, если из соотношения $x_n \to x$ следует

$$f(x_n) \to f(x).$$

Если M_1 есть совокупность вещественных чисел, то оператор называется *функционалом*.

Пример. Оператор Фредгольма в случае непрерывности ядра $K(s, t)$ есть непрерывный оператор, отображающий C в C.

Сфера. Будем называть *сферой* радиуса r с центром в точке a и обозначать через $S(a, r)$ совокупность точек $x \subset M$, для которых

$$\rho(x, a) < r.$$

Замкнутой сферой $\bar{S}(a, r)$ радиуса r назовем совокупность точек $x \subset M$, для которых

$$\rho(x, a) \leqslant r.$$

Заметим, что если $x \subset \bar{S}_1(a, r)$ и $x_1 \subset \bar{S}(a, r)$, то

$$\rho(x, x_1) \leqslant 2r. \tag{10}$$

В самом деле, из аксиомы треугольника следует

$$\rho(x, x_1) \leqslant \rho(x, a) + \rho(x_1, a) \leqslant r + r = 2r.$$

Точка a называется *предельной* для множества K, если всякая сфера $S(a, r)$ с центром в a содержит точки множества K. Множество, заключающее все свои предельные точки, называется *замкнутым*.

Множество $\bar{A}$ называется *замыканием* множества A, если оно получается путем прибавления к A всех его предельных точек. Если A замкнуто, то $\bar{A} = A$.

Сепарабельные пространства. Множество A называется *плотным* в пространстве M, если замыкание A совпадает с M.

Последовательность $x_1, x_2, \ldots, x_n, \ldots$ называется *счетной всюду плотной сетью*, если она плотна в M. Другими словами, каждый элемент $a \subset M$ есть предел для некоторой подпоследовательности $x_{k_1}, x_{k_2}, \ldots, x_{k_n}, \ldots$ Пространства, обладающие счетной всюду плотной сетью, называются *сепарабельными*.

Примеры. Числовая прямая сепарабельна. Счетную всюду плотную сеть на ней образуют рациональные точки. Точно так же n-мерные пространства сепарабельны. Пространства l_p сепарабельны. Каждая точка $x(x_1, x_2, \ldots, x_n, x_{n+1}, \ldots)$ из l_p есть предел последовательности точек $x^{(n)}(x_1, x_2, \ldots, x_n, 0, 0, \ldots) \subset l_p$. Точки $x^{(n)}$ являются в свою очередь предельными для точек $\xi^{(n)}(\xi_1, \xi_2, \ldots, \xi_n, 0, 0, \ldots)$ с рациональными ξ_i. Совокупность всех ξ образует счетную всюду плотную сеть на l_p. Аналогично можно доказать сепарабельность N.

Пространство C сепарабельно. Счетную всюду плотную сеть на нем образует совокупность полиномов с рациональными коэфициентами: в силу теоремы Вейерштрасса всякая непрерывная функция получается как предел равномерно сходящейся к ней последовательности полиномов; всякий же полином есть предел полиномов с рациональными коэфициентами.

Пространства L_p сепарабельны. Всякая неограниченная функция $y(t) \subset L_p$ есть предел в L_p ограниченных измеримых функций $y_N(t)$, где $y_N(t) = y(t)$, если $|y(t)| \leqslant N$, и $y_N(t) = 0$, если $|y(t)| > N$. Всякая же ограниченная измеримая функция есть предел в L_p непрерывных функций, а непрерывные функции суть пределы полиномов с рациональными коэфициентами (пределы в смысле равномерной сходимости, и тем более в смысле сходимости в L_p).

Аналогично доказывается сепарабельность пространства S.

Пространство m не сепарабельно. Рассмотрим совокупность X точек $x(\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots)$, где числа α_i равны нулю или единице. Точки из X образуют вершины „единичного куба" в m. Множество X имеет континуальную мощность. Расстояние между двумя различными точками из X равно единице (по крайней мере одна из их координат α_i равна единице в одной точке и нулю — в другой, разность этих координат равна единице). Построим сферы радиуса $\frac{1}{3}$ вокруг всех точек из X. Получается множество сфер без общих точек, имеющее континуальную мощность.

§ 2. Полнота.

При построении обычного анализа весьма существенную роль играет свойство *полноты* числовой прямой, которое эквивалентно критерию Коши сходимости числовых последовательностей. Естественно при построении общего анализа выделить те метрические пространства, где имеет место аналогичное свойство.

Определение. Последовательность x_n точек метрического пространства M называется *фундаментальной*, если для любого целого n существует число $\varepsilon_n \geqslant 0$ такое, что $\varepsilon_n \to 0$ при $n \to \infty$ и

$$\rho(x_n, x_{n+m}) \leqslant \varepsilon_n$$

при любом $m > 0$. (Иначе говоря, если расстояние между элементами последовательности с достаточно большими индексами может быть сделано сколь угодно малым.)

Замечания. 1. Всякая сходящаяся последовательность есть фундаментальная. В самом деле, пусть $x_n \to x$. Имеем

$$\rho(x_n, x_{n+m}) \leqslant \rho(x_n, x) + \rho(x_{n+m}, x).$$

Оба слагаемых в правой части при достаточно большом n могут быть сделаны сколь угодно малыми.

2. Если подпоследовательность x_{n_k} фундаментальной последовательности x_n сходится к x, то вся последовательность x_n сходится к x. В самом деле: при $n_k > n$, $\rho(x_n, x_{n_k}) < \varepsilon_n \to 0$. Далее,

$$\rho(x, x_n) \leqslant \rho(x, x_{n_k}) + \rho(x_n, x_{n_k}).$$

Но так как

$$\rho(x,\ x_{n_k}) \underset{n_k \to \infty}{\longrightarrow} 0,$$

то

$$\rho(x,\ x_n) \underset{n \to \infty}{\longrightarrow} 0.$$

3. Из определения фундаментальной последовательности следует, что расстояния от любого элемента фундаментальной последовательности a_n до других элементов a_k этой последовательности ограничены; так как $\rho(a, a_k) \leqslant \rho(a, a_n) + \rho(a_n, a_k)$, то и расстояния от произвольной точки a до элементов a_k фундаментальной последовательности ограничены.

Определение. Метрическое пространство M называется *полным*, если в нем всякая фундаментальная последовательность сходится.

Например, числовая прямая есть полное пространство. Можно доказать, что все приведенные на стр. 78—81 метрические пространства суть полные пространства. Для C и l_p это доказывается весьма просто.

Полнота пространства S. Пусть дана фундаментальная последовательность y_n в S. Из нее можно выбрать подпоследовательность $z_k = y_{n_k}$ такую, что

$$\rho(z_k,\ y_n) = \rho(y_{n_k},\ y_n) \leqslant \frac{1}{2^{k+2}}$$

для всех $n > n_k$. В частности

$$\rho(z_k,\ z_{k+1}) \leqslant \frac{1}{2^{k+2}}.$$

По определению метрики в S множество M_k точек $t \subset [0,\ 1]$, для которых

$$|z_{k+1}(t) - z_k(t)| > \frac{1}{2^{k+1}} > \frac{1}{2^{k+2}}, \tag{1}$$

имеет меру

$$\operatorname{mes} M_k \leqslant \frac{1}{2^{k+1}}.$$

Обозначим

$$S_i = \sum_{k=i}^{\infty} M_k, \quad T_i = [0,\ 1] - S_i. \tag{2}$$

Очевидно,

$$\operatorname{mes} S_i \leqslant \sum_{k=1}^{\infty} \frac{1}{2^{k+i}} = \frac{1}{2^i}; \quad \operatorname{mes} T_i \geqslant 1 - \frac{1}{2^i}; \quad \operatorname{mes} \sum_{i=1}^{\infty} T_i = 1. \tag{3}$$

В точках $t' \subset T_i$ при $k \geqslant i$ имеет место неравенство

$$|z_k(t') - z_{k+1}(t')| \leqslant \frac{1}{2^{k+1}}.$$

Отсюда

$$|z_{k+m}(t') - z_k(t')| \leqslant \sum_{i=0}^{m-1} |z_{k+i}(t') - z_{k+i+1}(t')| \leqslant \sum_{i=0}^{m-1} \frac{1}{2^{k+i+1}} < \frac{1}{2^k}. \tag{4}$$

Таким образом для всякого $t' \subset T_i$ последовательность $z_k(t')$, как фундаментальная числовая последовательность, сходится. Обозначим

$$z(t') = \lim_{k\to\infty} z_k(t').$$

$z(t)$ определена на всех точках, принадлежащих $\sum_{i=1}^{\infty} T_i$, т. е. [см. (3)] на всем отрезке [0, 1], исключая, быть может, множество меры 0; $z(t)$ есть элемент пространства S. Переходя к пределу в (4) по $m\to\infty$, имеем

$$|z(t') - z_k(t')| \leqslant \frac{1}{2^k} \tag{5}$$

всюду на множестве T_i (при $i \leqslant k$), в частности на T_k; так как

$$\operatorname{mes} T_k = 1 - \frac{1}{2^{k+1}} > 1 - \frac{1}{2^k},$$

то по определению метрики в S

$$\rho(z, z_k) \leqslant \frac{1}{2^k}.$$

Итак, $z_k \to z$. Так как z_k есть часть фундаментальной последовательности y_n, то имеем также (см. выше, стр. 82, замечание 2)

$$y_n \to z.$$

Тем самым полнота пространства S доказана.

Полнота пространства L_p. Пусть имеем фундаментальную последовательность y_n в L_p

$$[\rho(y_n, y_{n+m})]^p = \int_0^1 |y_n - y_{n+m}|^p\, dt \leqslant \varepsilon_n, \quad \varepsilon_n \to 0. \tag{6}$$

Расстояния y_n от $y \equiv 0$ ограничены (см. выше, стр. 83, замечание 3), т. е. существует константа K, для которой

$$\int_0^1 |y_n|^p\, dt \leqslant K, \quad n = 1, 2, \ldots \tag{7}$$

Выражение $|y_n(t) - y_{n+m}(t)|$ в силу неравенства (6) может превышать $\sqrt[2p]{\varepsilon_n}$ лишь на множество меры, не большей $\sqrt{\varepsilon_n}$. Следовательно,

$$\rho_S(y_n, y_{n+m}) \leqslant \sqrt[2p]{\varepsilon_n}\ ^{1)},$$

и y_n есть последовательность, фундаментальная также в смысле метрики в S. Повторив предыдущие рассуждения, сохраняя прежний смысл обозначений z_k, S_i, T_i, мы можем выделить из y_n подпоследовательность $z_k = y_{n_k}$, для которой удовлетворяются неравенства (1), (3) и (4) и, кроме того, в силу (6)

$$\int_0^1 |z_k - z_{k+l}|^p\, dt = \int_0^1 |y_{n_k} - y_{n_{k+l}}|^p\, dt \leqslant \varepsilon_{n_k} = \varepsilon'_k \quad (\varepsilon'_k \to 0). \tag{8}$$

---

[1] Символ $\rho_S(y_n, y_{n+m})$ указывает, что расстояние между y_n и y_{n+m} берется в смысле метрики в S.

Последовательность $z_k(t)$ стремится по мере к некоторой функции $z(t)$. При этом в силу неравенства (5) z_k равномерно сходится к z на любом T_i. Так как $\operatorname{mes} T_i \underset{i\to\infty}{\to} 1$, то

$$\int_0^1 |z|^p\,dt = \lim_{i\to\infty}\int_{T_i}|z|^p\,dt = \lim_{i\to\infty}\left[\lim_{k\to\infty}\int_{T_i}|z_k|^p\,dt\right] \leqslant \overline{\lim_{k\to\infty}}\int_0^1 |z_k|^p\,dt.$$

Поэтому [ср. (7)] $\int_0^1 |z|^p\,dt \leqslant K$ и $z(t) \subset L_p$.

Докажем сходимость z_n к z в смысле сходимости в L_p. Полагая $i \leqslant k$, имеем [ср. (3) и (5)]

$$\int_0^1 |z-z_k|^p\,dt = \int_{T_i}|z-z_k|^p\,dt + \int_{S_i}|z-z_k|^p\,dt \leqslant$$
$$\leqslant \frac{1}{2^{kp}} + \int_{S_i}|z-z_k|^p\,dt. \tag{9}$$

В силу неравенства Минковского и (8), полагая $n \leqslant k$, имеем

$$\int_{S_i}|z-z_k|^p\,dt \leqslant \int_{S_i}|z-z_n|^p\,dt + \int_{S_i}|z_n-z_k|^p\,dt \leqslant$$
$$\leqslant \int_{S_i}|z-z_n|^p\,dt + \int_0^1 |z_n-z_k|^p\,dt \leqslant \int_{S_i}|z-z_n|^p\,dt + \varepsilon_n'. \tag{10}$$

При фиксированном n и $i\to\infty$ имеем $\operatorname{mes} S_i \to 0$, и поэтому

$$\int_{S_i}|z-z_n|^p\,dt \to 0. \tag{11}$$

Выбрав произвольное $\eta > 0$, можно добиться при достаточно большом n, чтобы $\varepsilon_n' < \frac{\eta}{3}$; далее, зафиксировав n, можно выбрать i настолько большим, чтобы [см. (11)] $\int_{S_i}|z-z_n|^p\,dt < \frac{\eta}{3}$. И, наконец, выбираем k, большее, чем n и i, таким, чтобы $\frac{1}{2^{kp}} < \frac{\eta}{3}$. Из (9) и (10) следует, что

$$\int_0^1 |z-z_k|^p\,dt < \frac{\eta}{3}+\frac{\eta}{3}+\frac{\eta}{3} = \eta,$$

т. е. $z_k \to z$ в смысле сходимости в L_p.

Так как последовательность z_k есть часть фундаментальной последовательности y_n, то и

$$y_n \to z$$

в смысле сходимости в L_p.

Итак, полнота пространства L_p доказана.

Применение. Теорема Фишера-Рисса. Мы будем называть два метрических пространства M и M_1 *изометричными*, если можно установитьтакое взаимно-однозначное соответствие между их точками, при котором расстояние между любой парой точек в M равно расстоянию между соответственными точками в M_1 (и обратно).

Теорема 1 (Фишера-Рисса). *L_2 изометрично l_2 (гильбертову пространству).*

Рассмотрим в L_2 систему ортогональных функций $\varphi_1, \varphi_2, \varphi_3, \ldots, \varphi_n, \ldots$, т. е. таких функций, что

$$\int_0^1 \varphi_i \varphi_j \, dt = \begin{cases} 0 \text{ при } i \neq j, \\ 1 \text{ при } i = j. \end{cases}$$

Система φ_n называется *полной*, если всякая функция $y \subset L_2$ (с интегрируемым квадратом) представима (в смысле сходимости в среднем) своим *рядом Фурье* по φ_n,

$$\sum_{n=1}^{\infty} a_n \varphi_n(t),$$

где

$$a_n = \int_0^1 y(t) \, \varphi_n(t) \, dt$$

— *коэфициенты Фурье* функции $y(t)$. Условием полноты системы φ_n является, как нетрудно показать, выполнение *равенства Парсеваля*

$$\int_0^1 y^2 \, dt = \sum_{i=1}^{\infty} a_n^2. \tag{12}$$

Примерами полных ортогональных систем могут служить система тригонометрических функций

$$1, \ \sqrt{2} \sin \pi t, \ \sqrt{2} \cos \pi t, \ldots, \ \sqrt{2} \sin n\pi t, \ \sqrt{2} \cos n\pi t, \ldots,$$

система полиномов Лежандра и т. п.

Отнесем каждой $y(t) \subset L_2$ точку $a(a_1, a_2, \ldots, a_n, \ldots)$ из l_2, где a_n — коэфициенты Фурье функции $y(t)$. Пусть мы имеем функции $y(t)$ и $y_1(t)$ из L_2 с коэфициентами Фурье, соответственно, (a_n), (b_n); $y_1(t)$ отвечает точка $a(a_n) \subset l_2$, а $y_1(t)$ — точка $b(b_n) \subset l_2$.

Коэфициенты Фурье для $y(t) - y_1(t)$ суть $a_n - b_n$. Поэтому равенство Парсеваля для $y - y_1$ дает

$$\int_0^1 (y - y_1)^2 \, dt = \sum_{n=0}^{\infty} (a_n - b_n)^2. \tag{13}$$

Формула (13) показывает, что расстояние между $y(t)$ и $y_1(t)$ в L_2 равно расстоянию между a и b в l_2. Итак, L_2 изометрично некоторой части l_2. Если мы еще докажем, что при нашем соответствии всякому элементу $a \subset l_2$ отвечает некоторый элемент $y(t) \subset L_2$, то тем самым будет доказано, что L_2 изометрично l_2.

Пусть $a(a_n) \subset l_2$, т. е. $\sum_{n=1}^{\infty} a_n^2$ сходится. Рассмотрим функции

$$y_n(t) = \sum_{i=1}^{n} a_i \varphi_i(t).$$

Имеем в силу (12)

$$\int_0^1 [y_n(t) - y_{n+m}(t)]^2 \, dt = \int_0^1 \Big(\sum_{i=n+1}^{n+m} a_i \varphi_i(t) \Big)^2 dt = \sum_{i=n+1}^{n+m} a_i^2.$$

Вследствие сходимости ряда $\sum_{i=1}^{\infty} a_i^2$ имеем при любом $m > 0$

$$\sum_{i=n+1}^{n+m} a_i^2 \leqslant \varepsilon_n, \quad \varepsilon_n \to 0 \text{ при } n \to \infty.$$

Таким образом последовательность $y_n(t)$ есть фундаментальная в L_2 и в силу полноты L_2 сходится в среднем к некоторой функции $y(t) \subset L_2$. Но тогда

$$\int_0^1 y(t) \varphi_m(t) \, dt = \lim_{n \to \infty} \int_0^1 y_n(t) \varphi_m(t) \, dt \quad (m = 1, 2, \ldots).$$

В самом деле, в силу неравенства Шварца

$$\left| \int_0^1 (y - y_n) \varphi_m \, dt \right| \leqslant \sqrt{\int_0^1 \varphi_m^2 \, dt \int_0^1 (y - y_n)^2 \, dt} = \rho(y, y_n) \to 0.$$

Но для всех достаточно больших n

$$\int_0^1 y_n \varphi_m \, dt = \int_0^1 \sum_{k=1}^{n} a_k \varphi_k \varphi_m \, dt = a_m.$$

Таким образом

$$\int_0^1 y \varphi_m \, dt = a_m$$

и $y(t)$ соответствует точка $a(a_m)$ при нашем отображении L_2 на l_2.

Итак, изометричность L_2 и l_2 доказана.

Свойства полных пространств.

Теорема 2. *Пусть дана в полном пространстве M последовательность вложенных друг в друга сфер $\bar{S}(a_n, r_n)$ таких, что $r_n \to 0$. Существует точка a, общая всем $\bar{S}(a_n, r_n)$.*

Эта теорема является обобщением аналогичных свойств числовой прямой.

Для доказательства заметим, что при $p > 0$ сфера $\bar{S}(a_{n+p}, r_{n+p})$ с ее центром a_{n+p} входит в n-ю сферу $\bar{S}(a_n, r_n)$, откуда

$$\rho(a_{n+p}, a_n) \leqslant r_n.$$

Так как $r_n \to 0$, то последовательность центров a_n есть фундаментальная последовательность; ввиду полноты пространства M последовательность a_n имеет предельный элемент a: $a = \lim_{n \to \infty} a_n$. Так как все a_{n+p} входят при $p \geqslant 0$ в $\bar{S}(a_n, r_n)$, то и предел их a входит в $\bar{S}(a_n, r_n)$. Тем самым теорема доказана.

Докажем еще одно важное свойство полных пространств.

Мы будем говорить, что множество K *нигде не плотно* на M, если всякая сфера S на M содержит сферу S_1, не заключающую точек из K. (Иначе говоря, всякая сфера S на M заключает точки, отличные от $\overline{K}$.)

Теорема 3. *Полное пространство M не может быть суммой счетной последовательности множеств K_n, нигде не плотных на M.*

Построим для доказательства последовательность замкнутых сфер $\overline{S}_n$ следующим образом:

$\overline{S}_1$ есть произвольная замкнутая сфера радиуса $\leqslant 1$, не заключающая точек K_1. Если $\overline{S}_{n-1}$ уже построена, то в качестве $\overline{S}_n$ мы выберем заключенную в $\overline{S}_{n-1}$ замкнутую сферу, не содержащую точек K_n и радиуса $\leqslant \frac{1}{n}$. В силу того, что K_n нигде не плотно, выбрать в $\overline{S}_{n-1}$ сферу, не содержащую точек K_n, возможно (если радиус такой сферы оказался бы больше $\frac{1}{n}$, то мы взяли бы в качестве $\overline{S}_n$ концентрическую с ней сферу радиуса $\leqslant \frac{1}{n}$).

В силу теоремы 2 последовательность сфер $\overline{S}_n$ должна была бы заключать точку a, общую всем $\overline{S}_n$. По определению $\overline{S}_n$, a не входило бы ни в одно K_n. Следовательно, $\sum\limits_{i=1}^{\infty} K_n$ не может дать всего пространства M.

Пополнение пространств. Неполные пространства можно „пополнить" — сделать полными — путем введения новых идеальных элементов (подобно тому, как мы „пополняли" совокупность рациональных чисел до всей числовой прямой). Пусть a_n есть фундаментальная последовательность, не имеющая на M предельного элемента. Отнесем этой последовательности „идеальный" элемент a, который мы будем называть „пределом" последовательности a_n. Дополним пространство M всеми такими пределами. Если a и b — два присоединенных к M элемента — „пределы" последовательностей a_n и b_n из M, то установим их расстояние следующим образом:

$$\rho(a, b) = \lim_{n \to \infty} \rho(a_n, b_n).$$

Аналогично, если $a \subset M$, а $b = \lim\limits_{n \to \infty} b_n$ — присоединенный элемент, то $\rho(a, b) = \lim\limits_{n \to \infty} \rho(a, b_n)$. Если $\rho(a, b) = 0$, то a и b будем считать идентичными. Легко доказать, что с присоединением этих предельных элементов пространство M переходит в новое пространство M_1, попрежнему метрическое, но на этот раз уже полное.

Пример. Совокупность L'_p *непрерывных* функций $y(t)$, $0 \leqslant t \leqslant 1$, где

$$\rho(y, z) = \left(\int_0^1 |y - z|^p \, dt\right)^{\frac{1}{p}},$$

есть неполное пространство: последовательность непрерывных функций y_n, сходящаяся в среднем (p-й степени) к разрывной функции, есть последовательность, фундаментальная на L'_p, но не имеет на нем предела. Оказывается, что, пополнив L'_p, мы получим все пространство L_p.

§ 3. Компактность.

В обычной теории непрерывных функций основную роль играет свойство *компактности* отрезка числовой прямой (принцип Больцано-Вейерштрасса): из всякой ограниченной последовательности чисел можно выбрать сходящуюся числовую подпоследовательность. На этом свойстве построена теория абсолютного экстремума (теорема Вейерштрасса об ограниченности непрерывных на отрезке функций и о достижении ими своего максимума и минимума).

Определим компактное множество в любом абстрактном пространстве. Множество X пространства M называется *компактным*, если из любой последовательности x_n его элементов x можно выбрать сходящуюся подпоследовательность x_{n_i}. Если, сверх того, предельный элемент каждой такой последовательности принадлежит самому X (X замкнуто), то множество X называется *компактным в себе*.

Примеры. Отрезок прямой компактен в себе. Замкнутая сфера в n-мерном пространстве компактна в себе. Замкнутая сфера в C некомпактна: совокупность $y(t) \subset C$, для которых $|y(t)| \leqslant 1$ $(0 \leqslant t \leqslant 1)$, образует в C сферу радиуса 1 с центром в функции $y \equiv 0$. Последовательность $\sin n\pi x$, $n = 1, 2, 3, \ldots$ принадлежит этой сфере. Но ни эта последовательность, ни любая ее подпоследовательность не является равномерно сходящейся.

Для функционалов, определенных на компактных в себе множествах, имеет место

Теорема 1 (Вейерштрасса). *Пусть $f(x)$ — непрерывный функционал, определенный на компактном в себе множестве X. Тогда*

1. *Функционал $f(x)$ ограничен снизу (и сверху).*
2. *Функционал $f(x)$ достигает на X своей нижней (и верхней) грани.*

Доказательство. 1. Предположим, что $f(x)$ не ограничен снизу на X Тогда существует последовательность $x_n \subset X$, для которой

$$f(x_n) \to -\infty. \tag{1}$$

С другой стороны, в силу того, что X компактно в себе, из последовательности x_n можно выбрать подпоследовательность x_{n_i}, сходящуюся к некоторому $x \subset X$. Имеем вследствие непрерывности f

$$f(x_{n_i}) \to f(x). \tag{2}$$

С другой стороны, в силу (1)

$$f(x_{n_i}) \to -\infty.$$

Полученное противоречие и доказывает свойство 1.

2. Если c есть нижняя грань $f(x)$ на X, то существует на X последовательность x'_n, для которой

$$f(x'_n) \to c.$$

Из x'_n можно выбрать подпоследовательность x'_{n_i}, сходящуюся к некоторому элементу $x' \subset X$. Имеем

$$f(x') = \lim_{i \to \infty} f(x'_{n_i}) = c,$$

и свойство 2 доказано.

Теорема Вейерштрасса распространяется и на случай так называемых полунепрерывных функционалов. Функционал $f(x)$ называется *полунепрерывным снизу* (*сверху*), если из условия $x_n \to x$ следует, что

$$\underline{\lim}\, f(x_n) \geqslant f(x) \qquad [\overline{\lim}\, f(x_n) \leqslant f(x)].$$

Приведенное доказательство теорем Вейерштрасса переносится на случай полунепрерывных функционалов: *функционал, полунепрерывный снизу* (*сверху*) *определенный на компактном в себе множестве, ограничен снизу* (*сверху*) *и достигает на нем своей нижней* (*верхней*) *грани.*

Примером полунепрерывного снизу функционала служит определенный в Q функционал $y(q)$, означающий длину кривой q. Если $q_n \to q$, то

$$\underline{\lim}\, y(q_n) \geqslant y(q).$$

В теории экстремумов функций конечного числа переменных теорема Вейерштрасса гарантирует существование минимума для функции, определенной на замкнутой сфере n-мерного пространства. При переходе к вариационному исчислению мы встречаемся с дополнительной трудностью — некомпактностью основных функциональных пространств и их сфер. Рассмотрим, например, функционал

$$Y(y) = \int_0^1 y^2\, dt,$$

определенный для всех непрерывных функций $y(t)$, $0 \leqslant t \leqslant 1$, принимающих граничные значения

$$y(0) = 0; \quad y(1) = 1.$$

Нижняя грань $Y(y)$ равна нулю, ибо для функции $y = t^n$ функционал

$$Y(y) = \int_0^1 t^{2n}\, dt = \frac{1}{2n+1},$$

и, значит, может быть сделан сколь угодно малым. С другой стороны, для любой непрерывной кривой y, $Y(y) > 0$. Следовательно, функционал $Y(y)$ не достигает своей нижней грани 0.

В этом заключается специфическая трудность доказательств существования в вариационном исчислении. Однако во многих вопросах удается доказать существование кривой, реализующей искомый минимум, используя свойства компактности тех классов функций, на которых функционал определен.

Например, совокупность L кривых, соединяющих пару точек на замкнутой поверхности и имеющих ограниченную длину, компактна в Q (теорема Гильберта, см. следующий параграф). Следовательно, длина кривой — полунепрерывный функционал в L — достигает на некоторой кривой q_0 в силу теоремы Вейерштрасса своего абсолютного минимума. Из общих теорем вариационного исчисления следует, что кривая q_0 есть геодезическая дуга. Мы получаем доказательство существования геодезической дуги, соединяющей любую пару точек замкнутой поверхности.

Здесь мы имеем простейший пример доказательств существования в вариационном исчислении, — вопрос, которому посвящена в настоящее время обширная литература.

Структура компактных множеств в метрических пространствах. Будем рассматривать множества X, расположенные в метрическом пространстве M. Мы назовем множество X_1 ε-*сетью* по отношению к множеству X, если каждой точке $x \subset X$ можно отнести точку $x_1 \subset X_1$ такую, что

$$\rho(x, x_1) \leqslant \varepsilon.$$

Например, совокупность целочисленных точек образует $\frac{1}{2}$-сеть на числовой прямой.

ε-сеть называется *конечной*, если она состоит из конечного числа точек.

Теорема 2 Гаусдорфа). *Для компактности множества X в метрическом пространстве M необходимо и, в случае полноты M, достаточно, чтобы для любого положительного числа ε можно было построить конечную ε-сеть для X.*

1. Условие необходимо. В самом деле, построим последовательность элементов x_n следующим образом: x_1 — произвольный элемент из X. Если элементы $x_1, x_2, \ldots, x_{n-1}$ уже выбраны, то в качестве x_n мы выберем произвольную точку из X, которая удалена от любого из x_i $(1 \leqslant i < n)$ на расстояние, большее чем ε. Если такого элемента нет, то последовательность $x_1, x_2, \ldots, x_{n-1}$ обрывается и образует ε-сеть.

Последовательность x_n оборвется на некотором конечном элементе. В противном случае мы имели бы бесконечную последовательность x_n, для которой $\rho(x_n, x_m) > \varepsilon$, если $n \neq m$. Но тогда ни последовательность x_n, ни любая ее подпоследовательность не могли бы сходиться, что, однако, противоречит свойству компактности множества X, содержащего последовательность x_n. Итак, необходимость доказана.

2. Условие достаточно. Положим, что пространство M — полное и что множество X удовлетворяет условиям теоремы. Докажем, что из любой последовательности $T(x_1, x_2, \ldots, x_n, \ldots)$ точек из X можно выбрать сходящуюся подпоследовательность.

Выбрав последовательность положительных чисел ε_n, стремящуюся к нулю, мы по условию теоремы для каждого числа ε_n построим конечную ε_n-сеть

$$(a_1^{(n)}, a_2^{(n)}, \ldots, a_{k_n}^{(n)}).$$

Построим теперь подпоследовательности $T_1, T_2, \ldots, T_n, \ldots$ последовательности T следующим образом. Все точки X, и в частности точки последовательности T, заключены в конечном числе сфер

$$S(a_1^{(1)}, \varepsilon_1), \ S(a_2^{(1)}, \varepsilon_1), \ \ldots, \ S(a_{k_1}^{(1)}, \varepsilon_1)$$

радиуса ε_1 с центрами в точках ε_1-сети. По крайней мере одна из этих сфер, скажем $S(a_i^{(1)}, \varepsilon_1)$, заключает бесконечную подпоследовательность T_1 точек последовательности T. Пусть теперь уже выбраны подпоследовательности

$$T_1, T_2, \ldots, T_{n-1}.$$

Замечаем, что все точки подпоследовательности $T_{n-1} \subset X$ заключены в конечном числе сфер $S(a_i^{(n)}, \varepsilon_n)$, $i = 1, 2, \ldots, k_n$, с центрами в точках ε_n-сети и радиусом, равным ε_n. Тогда в одну по крайней мере из этих сфер попадет бесконечная подпоследовательность T_n точек из T_{n-1}.

В силу построения все точки T_n лежат в сфере радиуса ε_n и, следовательно, если $a, b \subset T_n$, то

$$\rho(a, b) \leqslant 2\varepsilon_n.$$

Составим теперь диагональным методом последовательность T_ω

$$a_1, \ a_2, \ \ldots, \ a_n, \ \ldots,$$

где

$$a_n \subset T_n.$$

При $m > 0$ имеем

$$T_{n+m} \subset T_n.$$

Отсюда

$$a_{n+m} \subset T_{n+m} \subset T_n, \quad \rho(a_n, a_{n+m}) \leqslant 2\varepsilon_n.$$

Последовательность T_ω есть фундаментальная последовательность; в силу полноты M она сходится.

Итак, мы выделили из последовательности T сходящуюся подпоследовательность T_ω.

В некоторых случаях полезно пользоваться следующим обобщением теоремы 2.

Теорема 3. *Для того чтобы множество X в полном метрическом пространстве M было компактным, достаточно, чтобы для любого $\varepsilon > 0$ существовала компактная ε-сеть относительно X.*

В самом деле, пусть удовлетворяются условия теоремы. При любом $\varepsilon > 0$ существует для X компактная $\frac{\varepsilon}{2}$-сеть X_1. Применяя к X_1 теорему 2, получаем, что для X_1 существует конечная $\frac{\varepsilon}{2}$-сеть X_2. Но X_2 есть ε-сеть для X. В самом деле, для любого $x \subset X$ существует точка $x_1 \subset X_1$ такая, что $\rho(x, x_1) \leqslant \frac{\varepsilon}{2}$. В свою очередь для x_1 существует такая точка $x_2 \subset X_2$, что $\rho(x_1, x_2) \leqslant \frac{\varepsilon}{2}$. Очевидно $\rho(x, x_2) \leqslant \rho(x, x_1) + \rho(x_1, x_2) \leqslant \varepsilon$ и X_2 есть конечная ε-сеть для X. Итак, при любом $\varepsilon > 0$ существует для X конечная ε-сеть. В силу теоремы 2 X компактно.

§ 4. Применения. Теоремы Арцела, Гильберта, Колмогорова, Бохнера.

Компактность в C. Назовем совокупность Y функций $y(t)$, $0 \leqslant t \leqslant 1$, *равномерно ограниченной*, если для всех $y \subset Y$ и для любых $t \subset [0, 1]$ имеем

$$|y(t)| \leqslant m \tag{1}$$

(m — константа).

Назовем совокупность Y функций $y(t)$, $0 \leqslant t \leqslant 1$, *равностепенно непрерывной*, если удовлетворяется следующее условие: всякому числу $h > 0$ отвечает число $\varepsilon = \varepsilon(h) \geqslant 0$ такое, что $\varepsilon(h) \to 0$ при $h \to 0$ и

$$|y(t) - y(t_1)| < \varepsilon, \tag{2}$$

коль скоро $y \subset Y$; $t, t_1 \subset [0, 1]$; $|t - t_1| \leqslant h$, причем число $\varepsilon = \varepsilon(h)$ зависит только от h и *не зависит* от выбора функций y и интервала (t, t_1).

Теорема 1 (Арцела). *Для того чтобы совокупность Y непрерывных на $[0, 1]$ функций $y(t)$ была компактна в C (компактна в смысле равномерной сходимости), необходимо и достаточно, чтобы Y было равномерно ограниченной и равностепенно непрерывной системой функций.*

1. Условие необходимо. Пусть Y компактна в C. В силу теоремы 2, § 3, для всякого ε существует на Y конечная $\frac{\varepsilon}{3}$-сеть $(y_1, y_2, \ldots, y_n)$. Так как $y_i(t)$ непрерывны, то каждая из них ограничена:

$$|y_i(t)| \leqslant m_i \quad (0 \leqslant t \leqslant 1) \tag{3}$$

(m_i — константы). Пусть m — наибольшее из чисел m_i.

С другой стороны, для каждого $y(t) \subset Y$ существует такое $y_i(t)$ из нашей $\frac{\varepsilon}{3}$-сети, что

$$\rho(y, y_i) = \max_{0 \leqslant t \leqslant 1} |y(t) - y_i(t)| \leqslant \frac{\varepsilon}{3}. \tag{4}$$

Из (3) и (4) следует, что

$$|y(t)| \leqslant |y_i(t)| + |y(t) - y_i(t)| \leqslant m + \frac{\varepsilon}{3} = m' \tag{5}$$

(m' — константа).

Итак, вся система Y равномерно ограничена.

Для каждой из наших n функций $y_i(t)$ можно в силу их непрерывности утверждать: для любого $h > 0$ существует такое $\varepsilon_i(h)$, $\varepsilon_i(h) \to 0$ при $h \to 0$, что

$$|y_i(t) - y_i(t_1)| \leqslant \varepsilon_i(h), \tag{6}$$

коль скоро $|t - t_1| \leqslant h$; $t, t_1 \subset [0, 1]$.

Обозначим через $\varepsilon_0(h)$ наибольшее из чисел $\varepsilon_i(h)$, $i = 1, 2, \ldots, n$. Выберем h таким, чтобы $\varepsilon_0(h) \leqslant \frac{\varepsilon}{3}$.

Для произвольной $y \subset Y$ при $|t - t_1| \leqslant h$ и $t, t_1 \subset [0, 1]$ имеем [в силу неравенств (4) и (6)]

$$|y(t) - y(t_1)| \leqslant |y(t) - y_i(t)| + |y_i(t) - y_i(t_1)| + |y_i(t_1) - y(t_1)| \leqslant$$
$$\leqslant \frac{\varepsilon}{3} + \varepsilon_0(h) + \frac{\varepsilon}{3} \leqslant \varepsilon.$$

Итак, Y есть равностепенно непрерывная система функций.

2. Условие достаточно. Пусть, в самом деле, удовлетворяются неравенства (1) и (2) для всех $y \subset Y$. Выберем произвольное $\varepsilon_0 > 0$. Разобьем отрезок $[0, 1]$ на n частей $\left[\frac{i}{n}, \frac{i+1}{n}\right]$ $(i = 0, 1, \ldots, n-1)$. При достаточно большом n

$$\varepsilon\left(\frac{1}{n}\right) \leqslant \varepsilon_0$$

и для всех $y \subset Y$

$$|y(t_2) - y(t_1)| \leqslant \varepsilon_0, \quad \text{если} \quad |t_2 - t_1| \leqslant \frac{1}{n}; \quad t_2, t_1 \subset [0, 1], \tag{7}$$

в частности, если

$$t_2, t_1 \subset \left[\frac{i}{n}, \frac{i+1}{n}\right].$$

Отнесем каждой функции $y(t) \subset Y$ полигональную функцию $y_n(t)$ такую, что

$$y_n\left(\frac{i}{n}\right) = y\left(\frac{i}{n}\right) \quad (i = 0,\ 1,\ \ldots,\ n),$$

а в интервалах $\left[\frac{i}{n}, \frac{i+1}{n}\right]$ y_n линейна; $y = y_n(t)$ есть полигон с n сторонами, вписанный в кривую $y = y(t)$.

Совокупность Y_n полигонов $y_n(t)$ для $y(t) \subset Y$ образует ε_0-сеть на Y. В самом деле, для

$$t \subset \left[\frac{i}{n}, \frac{i+1}{n}\right] \quad (i = 0,\ 1,\ \ldots,\ n-1)$$

$y_n(t)$ вследствие линейности заключено между $y_n\left(\frac{i}{n}\right) = y\left(\frac{i}{n}\right)$ и $y_n\left(\frac{i+1}{n}\right) = y\left(\frac{i+1}{n}\right)$. Следовательно, $|y(t) - y_n(t)|$ заключено между $\left|y(t) - y\left(\frac{i}{n}\right)\right|$ и $\left|y(t) - y\left(\frac{i+1}{n}\right)\right|$, и в силу неравенства (7)

$$|y(t) - y_n(t)| \leqslant \varepsilon_0.$$

Итак, Y_n образует ε_0-сеть на Y. Но $Y_n(t)$ образует ограниченное множество в $(n+1)$-мерном пространстве (каждый полигон $y_n(t)$ определяется $(n+1)$-й ординатами его вершин). Следовательно, Y_n компактно.

Итак, для любого $\varepsilon_0 > 0$ можно построить компактную ε_0-сеть относительно Y. В силу теоремы 3, § 3 Y компактно.

Компактность в Q.

Теорема 2 (Гильберта). *Совокупность Q_1 спрямляемых кривых ограниченной длины, расположенных в ограниченной части плоскости, компактна (в смысле сходимости в Q).*

Пусть длины кривых $q \subset Q_1$ не превосходят l. Разобьем каждую кривую $q \subset Q_1$ на n дуг равной длины и соединим точки деления отрезками. Получим полигон q_n. Каждая из n дуг кривой q и соответственных сторон полигона q_n имеет длину, не превосходящую $\frac{l}{n}$. Расстояние между точками такой дуги и точками стягивающей ее хорды — стороны q_n — не превосходит $\frac{2l}{n}$. Введем на q и q_n такие параметрические представления, чтобы вершинам полигона отвечали в обоих представлениях числа $\frac{i}{n}$ $(i = 0,\ 1,\ 2,\ \ldots,\ n)$ и чтобы когда t пробегает интервал $\left(\frac{i}{n}, \frac{i+1}{n}\right)$, мы получали одну из дуг в q и соответствующую сторону в q_n.

Расстояние между точками q и q_n, отвечающими одинаковому значению параметра, не превосходит $\frac{2l}{n}$. По определению $\rho(q, q_n)$ в Q имеем (см. § 1, пример 9)

$$\rho(q, q_n) \leqslant \frac{2l}{n}.$$

Совокупность $Q_1^{(n)}$ полигонов q_n образует $\frac{2l}{n}$-сеть на Q_1. Каждый полигон q_n определяется $2(n+1)$ координатами его $n+1$ вершин. В силу принципа Больцано-Вейерштрасса Q_1 компактно. По теореме 3, § 3 компактно и Q.

Компактность в L_p. Мы будем в дальнейшем считать, что функция $y(t) \subset L_p$, определенная на $[0, 1]$, равна нулю вне этого отрезка.

Введем обозначение

$$y_h(t) = \frac{1}{2h} \int\limits_{t-h}^{t+h} y(s)\, ds.$$

В $y_h(t)$ значения функции $y(t)$ заменяются средними значениями этой функции на интервалах длины $2h$.

Воспользуемся неравенством Гельдера [1])

$$\left| \int\limits_a^b uv\, dt \right| \leqslant \left\{ \int\limits_a^b |u|^p dt \right\}^{\frac{1}{p}} \left\{ \int\limits_a^b |v|^{\frac{p}{p-1}} dt \right\}^{\frac{p-1}{p}} \quad (b>a). \tag{8}$$

Отсюда

$$\left| \int\limits_a^b y\, dt \right| = \left| \int\limits_a^b y \cdot 1 \cdot dt \right| \leqslant \left\{ \int\limits_a^b |y|^p dt \right\}^{\frac{1}{p}} \left\{ \int\limits_a^b 1 \cdot dt \right\}^{\frac{p-1}{p}} =$$

$$= (b-a)^{\frac{p-1}{p}} \left\{ \int\limits_a^b |y|^p dt \right\}^{\frac{1}{p}}. \tag{9}$$

Как следствие из (9) получаем

$$\rho_{L_p}(y_h, z_h) \leqslant \rho_{L_p}(y, z). \tag{10}$$

В самом деле

$$\int\limits_0^1 |y_h|^p dt = \frac{1}{(2h)^p} \int\limits_0^1 \left| \int\limits_{t-h}^{t+h} y\, dt_1 \right|^p dt \leqslant \frac{1}{(2h)^p} \int\limits_0^1 \left[(2h)^{\frac{p-1}{p}} \left\{ \int\limits_{t-h}^{t+h} |y|^p dt_1 \right\}^{\frac{1}{p}} \right]^p dt =$$

$$= \frac{1}{2h} \int\limits_0^1 \int\limits_{t-h}^{t+h} |y|^p dt_1\, dt = \frac{1}{2h} \int\limits_{-h}^{h} dt_1 \int\limits_{t_1}^{1+t_1} |y|^p dt \leqslant \int\limits_0^1 |y|^p dt. \tag{11}$$

(Последнее неравенство вытекает из того, что по условию, вне $[0, 1]$ $y=0$, откуда

$$\int\limits_{t_1}^{1+t_1} |y|^p dt \leqslant \int\limits_0^1 |y|^p dt.)$$

---

[1]) Неравенство Гельдера

$$\left\{ \sum_{i=1}^n |u_i|^p \right\}^{\frac{1}{p}} \left\{ \sum_{i=1}^n |v_i|^{\frac{p}{p-1}} \right\}^{\frac{p-1}{p}} \geqslant \left| \sum_{i=1}^n u_i v_i \right|$$

и аналогично для интегралов

$$\left\{ \int |u|^p dt \right\}^{\frac{1}{p}} \left\{ \int |v|^{\frac{p}{p-1}} dt \right\}^{\frac{p-1}{p}} \geqslant \left| \int uv\, dt \right|$$

можно доказать, например, определив обычным методом минимум разности его левой и правой частей.

Заменяя в формуле (11) y через $y-z$, мы и получаем

$$\int_0^1 |y_h - z_h|^p \, dt \leqslant \int_0^1 |y - z|^p \, dt,$$

т. е. неравенство (10).

Теорема 3 (А. Н. Колмогорова). *Для того чтобы система Y функций $y(t) \subset L_p$ была компактна в L_p, необходимо и достаточно, чтобы*

1) $\int_0^1 |y|^p \, dt \leqslant K^p$ *(K — константа)*;

2) *существовала функция $\eta(h) \geqslant 0$ для $h \geqslant 0$ такая, что $\eta(h) \to 0$ при $h \to 0$ и*

$$\rho(y, y_h) \leqslant \eta(h)$$

для всех $y \subset Y$[1]).

Заметим, что условие 1) означает

$$\rho(\theta, y) \leqslant K, \tag{12}$$

где θ есть функция, тождественно равная нулю.

1. Условия необходимы. В самом деле, в силу теоремы 2, § 3, для любого $\varepsilon > 0$ существует на Y конечная $\frac{\varepsilon}{3}$-сеть $(y_1, y_2, \ldots, y_n)$. Так как любую функцию в L_p можно аппроксимировать сколь угодно близкими к ней (в смысле расстояния в L_p) непрерывными функциями, то все y_i нашей $\frac{\varepsilon}{3}$-сети можно считать непрерывными. Для каждого $y \subset Y$ существует такая y_i $(1 \leqslant i \leqslant n)$, что

$$\rho(y, y_i) < \frac{\varepsilon}{3}. \tag{13}$$

Обозначая через m наибольшее из чисел $\rho(\theta, y_i)$, имеем [см. (13)]

$$\rho(\theta, y) \leqslant \rho(\theta, y_i) + \rho(y_i, y) \leqslant m + \frac{\varepsilon}{3}.$$

Таким образом условие 1) будет удовлетворено, если взять

$$K = m + \frac{\varepsilon}{3}.$$

Далее, для непрерывных функций, очевидно, $\rho(y_h, y) \to 0$, откуда для каждого y_i

$$\rho(y_i, (y_i)_h) = \eta_i(h) \to 0 \quad \text{при} \quad h \to 0. \tag{14}$$

Выберем h настолько малым, чтобы все n чисел $\eta_i(h)$ стали меньше $\frac{\varepsilon}{3}$. Имеем в силу (10), (13) и (14)

$$\rho(y, y_h) \leqslant \rho(y, y_i) + \rho(y_i, (y_i)_h) + \rho((y_i)_h, y_h) \leqslant$$
$$\leqslant 2\rho(y, y_i) + \rho(y_i, (y_i)_h) \leqslant \frac{2}{3}\varepsilon + \frac{\varepsilon}{3} = \varepsilon.$$

Обозначая $\varepsilon = \eta(h)$, мы видим, что условие 2) также выполнено.

[1]) Другую формулировку условий компактности L_p см. в статье В. В. Немыцкого „Метод неподвижных точек в анализе" в настоящем выпуске „Успехов математических наук", стр. 160.

2. Условия достаточны. Пусть условия 1) и 2) выполнены. В силу (9) и (12) при $0 \leqslant t_1 - t \leqslant h$ имеем для $y \subset Y$

$$|y_h(t_1) - y_h(t)| = \frac{1}{2h}\left|\int\limits_{t_1-h}^{t_1+h} y\,ds - \int\limits_{t-h}^{t+h} y\,ds\right| \leqslant$$

$$\leqslant \frac{1}{2h}\left(\left|\int\limits_{t-h}^{t_1-h} y\,ds\right| + \left|\int\limits_{t+h}^{t_1+h} y\,ds\right|\right) \leqslant$$

$$\leqslant \frac{(t_1-t)^{\frac{p-1}{p}}}{2h} \cdot 2\left\{\int\limits_0^1 |y|^p\,ds\right\}^{\frac{1}{p}} \leqslant \frac{K(t_1-t)^{\frac{p-1}{p}}}{h}. \quad (15)$$

Отсюда все y_h непрерывны.

Обозначим через Y_h совокупность функций y_h для $y \subset Y$. Из (15) мы видим, что Y есть равностепенно непрерывная система функций. Далее, отсюда же следует, что на всяком интервале длины $\leqslant h$ колебание $y_h \subset Y$ не превосходит $Kh^{-\frac{1}{p}}$. Если $\frac{1}{n} < h$, то мы видим, что на всем отрезке $[0, 1]$ колебание y_h не превосходит

$$K_1 = nK\,h^{-\frac{1}{p}}.$$

Из (10) и (12) мы имеем

$$\int\limits_0^1 |y_h|^p\,dt \leqslant K^p.$$

Значит всякая функция $y_h \subset Y$ должна в какой нибудь точке $x \subset [0, 1]$ принимать значение, не превосходящее K. А так как колебание y_h на $[0, 1]$ не превосходит K_1, то для любого $t \subset [0, 1]$

$$|y_h(t)| < K_2 = K + K_1.$$

Система Y_n равномерно ограничена.

По теореме Арцела система Y_h компактна в смысле равномерной сходимости, значит тем более она компактна в смысле сходимости в L_p.

По условию теоремы Y_h образует $\eta(h)$-сеть на Y. Итак, для любого $\varepsilon > 0$ мы при $\eta(h) < \varepsilon$ имеем для Y компактную ε-сеть. В силу теоремы 3, § 3 Y компактно.

Компактность систем почти периодических функций. Обозначим через K совокупность всех функций $y(t)$, $-\infty < t < \infty$, ограниченных на числовой прямой. В частности K включает все почти периодические по Бору функции[1]).

Определим для $y(t)$ и $y_1(t)$ из K расстояние так:

$$\rho(y, y_1) = \sup_{-\infty < t < \infty} |y(t) - y(t_1)|.$$

Сходимость в K означает, следовательно, равномерную сходимость на всей числовой прямой.

---

[1]) См., например, Г. Бор, Почти периодические функции, ГТТИ, 1933.

Теорема 4. *Для того чтобы совокупность Y почти периодических по Бору функций была компактна в смысле равномерной сходимости (т. е. сходимости в K), необходимо и достаточно, чтобы*

1) *Совокупность Y была равномерно ограничена.*

2) *Совокупность Y была равностепенно непрерывна.*

3) *Для каждого $\eta > 0$ существовало такое $l = l(\eta)$, чтобы всякий интервал $(a, a+l)$ числовой прямой содержал число s, являющееся одновременно η-периодом для всех функций $y \subset Y$:*

$$|y(t+s) - y(t)| \leqslant \eta, \quad -\infty < t < \infty. \tag{16}$$

1. Условия необходимы. Заметим, что всякая почти периодическая функция ограничена и равномерно непрерывна на числовой прямой. Поэтому доказательство необходимости условий 1) и 2) совершенно аналогично доказательству необходимости подобных условий в теореме Арцела.

Для доказательства необходимости условия 3) заметим, что для конечной системы почти периодических функций это условие удовлетворяется [1]).

Если Y компактна, то для любого $\eta > 0$ на Y существует конечная $\frac{\eta}{3}$-сеть, состоящая из функций $y_1(t)$, $y_2(t), \ldots, y_n(t)$ (принадлежащих к Y). Тогда, по сделанному только что замечанию, существует такое число $l > 0$, что всякий интервал $(a, a+l)$ содержит число s, являющееся $\frac{\eta}{3}$-периодом одновременно для всех $y_i(t)$, $i = 1, 2, \ldots, n$:

$$|y_i(t+s) - y_i(t)| \leqslant \frac{\eta}{3}, \quad i = 1, 2, \ldots, n, \quad -\infty < t < +\infty. \tag{17}$$

С другой стороны, так как $y_i(t)$ образуют $\frac{\eta}{3}$-сеть на Y, то для всякого $y \subset Y$ существует такое $y_i(t)$, $i = 1, 2, \ldots, n$, что

$$\rho(y, y_i) \leqslant \frac{\eta}{3}$$

или

$$|y(t) - y_i(t)| \leqslant \frac{\eta}{3}, \quad -\infty < t < \infty. \tag{18}$$

Из (17) и (18) следует, что

$$|y(t+s) - y(t)| \leqslant |y(t+s) - y_i(t+s)| + \\ + |y_i(t+s) - y_i(t)| + |y_i(t) - y(t)| \leqslant \frac{\eta}{3} + \frac{\eta}{3} + \frac{\eta}{3} = \eta \tag{19}$$

для $-\infty < t < \infty$.

Итак, s есть η-период для всякого y, и необходимость условия 3) доказана.

2. Условия достаточны. Предположим, что условия 1), 2) и 3) выполнены для системы Y почти периодических функций. Выберем произвольное $\eta > 0$ и пусть $l = l(\eta)$. Каждый интервал длины l содержит η-период для всех $y(t) \subset Y$.

---

[1]) См. Г. Бор, Почти периодические функции, ГТТИ, 1933, стр. 46—47.

Отнесем каждой $y(t) \subset Y$ следующую функцию $\bar{y}(t)$:

$$\bar{y}(t) = y(t) \quad \text{при} \quad -l \leqslant t \leqslant l,$$

$$\bar{y}(t) = y(t - r_n) \quad \text{при} \quad \begin{cases} nl < t \leqslant (n+1)\,l & \text{для} \quad n = 1, 2, 3, \ldots, \\ nl \leqslant t < (n+1)\,l & \text{для} \quad n = -2, -3, -4, \ldots, \end{cases}$$

где r_n есть общий η-период для всех $y \subset Y$, лежащий в интервале $(nl, (n+1)\,l)$. Заметим, что

$$t - r_n \subset [-l, l].$$

Обозначим совокупность всех $\bar{y}(t)$ через Y_η.

Если последовательность $\bar{y}_m(t) \subset Y_\eta$ сходится равномерно на отрезке $[-l, l]$, то (в силу определения этих функций) она сходится равномерно также на всей числовой прямой.

Совокупность всех $\bar{y} \subset Y_\eta$ на отрезке $[-l, l]$ удовлетворяет условиям теоремы Арцела. Поэтому Y_η компактно в смысле равномерной сходимости на этом отрезке. В силу сделанного только что замечания Y_η компактно в смысле равномерной сходимости на всей числовой прямой (Y_η компактно в K).

Для любого $y(t) \subset Y$ и соответственного $\bar{y}(t)$ имеем

$$y(t) - \bar{y}(t) = 0 \quad \text{при} \quad t \subset [-l, l],$$

$$y(t) - \bar{y}(t) = y(t) - y(t - r_n) \,\text{при} \begin{cases} nl < t \leqslant (n+1)\,l & \text{для} \quad n = 1, 2, \ldots, \\ nl \leqslant t < (n+1)\,l & \text{для} \quad n = -2, -3, \ldots \end{cases}$$

Так как r_n есть η-период для $y(t)$, то для любого t, $-\infty < t < \infty$,

$$|y(t) - \bar{y}(t)| < \eta.$$

Следовательно, компактное множество Y_η образует η-сеть на K. В силу теоремы 3, § 3, Y компактно, и достаточность условий 1)—3) доказана.

Определение почти периодической функции по Бохнеру. Пусть нам дана функция $y(t)$, $-\infty < t < \infty$. Обозначим

$$y_s(t) = y(t + s).$$

Непрерывная функция $y(t)$ называется *почти периодической по Бохнеру*, если совокупность $y_s(t)$, $-\infty < s < \infty$, компактна в смысле равномерной сходимости на числовой прямой.

Теорема 5. *Определения почти периодических функций по Бору и по Бохнеру эквивалентны.*

В самом деле, пусть $y(t)$ есть почти периодическая функция по Бору. Совокупность функций $y_s(t) = y(t+s)$ удовлетворяет, очевидно, всем трем условиям теоремы 4, откуда следует ее компактность.

Пусть теперь $y(t)$ почти периодична по Бохнеру. $y(t)$ ограничена на числовой прямой, $|y(t)| < K$, $-\infty < t < \infty$. В самом деле, если бы существовала последовательность чисел t_n, для которых $y(t_n) \to \infty$, то для последовательности функций $y_{t_n}(t) = y(t + t_n)$ мы имели бы $y_{t_n}(0) \to \infty$. Следовательно, ни эта последовательность функций, ни любая ее подпоследовательность не могли бы равномерно сходиться (вопреки определению почти периодичности по Бохнеру).

В силу теоремы 2,* § 3, компактная совокупность $\{y_s\}$ содержит конечную ε-сеть $y_{s_1}, y_{s_2}, \ldots, y_{s_n}$. Расположим числа s_i в возрастающем порядке: $s_1 < s_2 < \ldots < s_n$. Для любого $y_s(t) = y(t+s)$ существует такое s_i, $i = 1, 2, \ldots, n$, что

$$\rho(y_s, y_{s_i}) = \sup |y_s(t) - y_{s_i}(t)| \leqslant \varepsilon,$$

или

$$|y(t+s) - y(t+s_i)| \leqslant \varepsilon, \quad -\infty < t < \infty.$$

Полагая $t + s_i = t'$, имеем:

$$|y(t' + s - s_i) - y(t')| \leqslant \varepsilon, \quad -\infty < t' < \infty.$$

Тем самым мы показали, что для любого вещественного s одно из n чисел $s - s_i$ есть ε-период.

Отсюда вытекает, что каждый отрезок $[a, a + s_n - s_1]$ содержит ε-период функции y. В самом деле, полагая $a + s_n = s$, мы в силу предыдущего должны иметь одно из чисел $a + s_n - s_i$, $i = 1, 2, \ldots, n$, в качестве ε-периода функции $y(t)$. Но так как $s_1 \leqslant s_i \leqslant s_n$, то

$$a + s_n - s_1 \geqslant a + s_n - s_i \geqslant a.$$

Тем самым мы доказали, что $y(t)$ почти периодична в смысле Бора [1]).

§ 5. Линейные пространства.

n-мерное пространство, рассматриваемое как совокупность векторов, образует *линейную* систему: над векторами в n-мерном пространстве можно производить линейные операции — сложение и умножение на скаляр (вещественное число).

Мы называем *линейной системой* совокупность A элементов, над которыми можно производить линейные операции — сложение и умножение на скаляр (причем результаты этих операций снова принадлежат к A), удовлетворяющие следующим аксиомам (a, b, c — элементы пространства; λ, μ — скаляры):

1. Ассоциативность сложения:

$$(a + b) + c = a + (b + c).$$

2. Коммутативность сложения [2]):

$$a + b = b + a.$$

3. $$1 \cdot a = a.$$

4. Существует единственный нулевой элемент θ:

$$\theta = a - a = 0 \cdot a$$

(очевидно, $\lambda\theta = \lambda(a - a) = \lambda a - \lambda a = \theta$).

---

[1]) Преимущество определения Бохнера заключается в том, что оно допускает широкие обобщения: можно говорить о почти периодических по Бохнеру функциях $y(s)$, где s — элементы любой группы, а y — элементы любого метрического пространства. См. S. Bochner, Abstracte fastperiodische Funktionen, Acta Mathematica, т. 61, 1933, стр. 149—184; J. v. Neumann, Almost periodic functions in a group I, Transact. Amer. Math. Soc., т. 36, n° 3, 1934, стр. 445—492; S. Bochner and J. v. Neumann, то же, II, там же, т. 37, n° 1, 1935, стр. 21—50.

[2]) Аксиома 2 коммутативности сложения вытекает из остальных аксиом. В самом деле,

$$(a + b) - (b + a) = (a + b) - 1(b + a) = a + b - b - a = a + (b - b) - a = (a + \theta) - a = a - a = \theta.$$

5. Ассоциативность умножения:

$$\lambda(\mu a) = (\lambda\mu)\, a.$$

$$\left.\begin{array}{ll} 6. & \lambda(a+b) = \lambda a + \lambda b, \\ 7. & (\lambda+\mu)\, a = \lambda a + \mu a. \end{array}\right\} \text{ Два закона дистрибутивности.}$$

Если линейная система является в то же время абстрактным пространством, то она называется *линейным* или *векторным пространством* [1]).

Норма. Особо важным классом линейных пространств являются так называемые *пространства типа* B (Банаха). В этих пространствах определена норма элемента, по аналогии с нормой вектора. Каждому $a \subset B$ отвечает неотрицательное число $\|a\|$ (*норма* a), причем нормы удовлетворяют следующим аксиомам:

1. Если $\|a\| = 0$, то $a = \theta$.
2. $\|\lambda a\| = |\lambda|\,\|a\|$.
3. $\|a+b\| \leqslant \|a\| + \|b\|$ (аксиома треугольника).

Аксиомы 1—3 являются обобщением основных свойств норм векторов в евклидовом пространстве. Расстояние между элементами (точками) пространства B определяется формулой

$$\rho(a, b) = \|a - b\|.$$

Очевидно, при этом все аксиомы метрики удовлетворяются.

Примеры 1. Пространства $l_p^{(n)}$ суть пространства B, если определить в них линейные операции и нормы следующим образом:

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1+y_1, x_2+y_2, \ldots, x_n+y_n),$$

$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n),$$

$$\|(x_1, x_2, \ldots, x_n)\| = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}.$$

2. Аналогично обстоит дело и с l_p и m. Определяем:

$$(x_1, x_2, \ldots) + (y_1, y_2, \ldots) = (x_1+y_1, x_2+y_2, \ldots),$$

$$\lambda(x_1, x_2, \ldots) = (\lambda x_1, \lambda x_2, \ldots),$$

$$\|(x_1, x_2, \ldots)\| = \left\{\sum_{i=1}^{\infty} |x_i|^p\right\}^{\frac{1}{p}} \text{ в } l_p,$$

$$\|(x_1, x_2, \ldots)\| = \sup |x_i| \text{ в } m.$$

3. В функциональных пространствах L_p и C определяем для их элементов — функций $y(t)$ — сложение и умножение на скаляр как обычное сложение функций и умножение на скаляр. Норму функций $y(t)$ в L_p определяем так:

$$\|y\| = \left\{\int_0^1 |y|^p\, dt\right\}^{\frac{1}{p}}.$$

[1]) Теория линейных пространств, а также определенных на них линейных функционалов и операторов подробно изложена в книге Banach'a, Théorie des opérations linéaires, Warszawa 1932.

Норму функции $z(t)$ в C определяем так:

$$\|z\| = \max_{0 \leqslant t \leqslant 1} |z(t)|.$$

Пространства $l_p^{(n)}$, l_p, L_p, C дают нам примеры линейных пространств типа B.

Сферы. Сферы $S(a, r)$ и $\overline{S}(a, r)$ в линейном пространстве определяются как совокупности элементов, удовлетворяющих неравенствам

$$\|x - a\| < r, \quad \text{соответственно } \|x - a\| \leqslant r.$$

В частности, сферы $S(\theta, r)$ и $\overline{S}(\theta, r)$ определяются неравенствами

$$\|x\| < r, \quad \text{соответственно } \|x\| \leqslant r.$$

n-Мерные пространства. *n-мерные* линейные пространства суть те, все элементы которых выражаются линейно через конечное число их (например $x_1, x_2, \ldots, x_n$), причем n есть наименьшее число их элементов, через которые все элементы выражаются линейно.

Остальные пространства называются *бесконечномерными* (например l_p, L_p, C).

В n-мерных пространствах сферы компактны (принцип Больцано-Вейерштрасса). Наоборот, в бесконечномерных пространствах сферы некомпактны (см. стр. 82, конец § 2, и стр. 89, примеры).

Принцип Больцано-Вейерштрасса есть необходимое и достаточное условие конечномерности пространства.

Выпуклые множества. Определим (по аналогии с обычной векторной геометрией) *отрезок*, соединяющий точки a и b линейного пространства, как совокупность всех точек вида

$$ta + (1 - t)b, \quad 0 \leqslant t \leqslant 1.$$

Множество X называется *выпуклым*, если оно заключает всякий отрезок, соединяющий любые две его точки.

Примеры 1. Совокупность полиномов образует выпуклое множество в C и в L_p.

2. Любая сфера $\|x\| \leqslant r$ есть выпуклое множество.

В самом деле, пусть a и b входят в эту сферу: $\|a\| \leqslant r$, $\|b\| \leqslant r$. Тогда для любого t, $0 \leqslant t \leqslant 1$,

$$\|ta + (1 - t)b\| \leqslant \|ta\| + \|(1 - t)b\| = t\|a\| + (1 - t)\|b\| \leqslant$$
$$\leqslant tr + (1 - t)r = r.$$

Итак, весь отрезок, соединяющий a и b, входит в сферу $\|x\| \leqslant r$.

Можно доказать, что если потребовать удовлетворения аксиом 1 и 2 для нормы и выпуклости сферы $\|x\| \leqslant 1$, то удовлетворяется аксиома треугольника. В самом деле, тогда имеем

$$a + b = (\|a\| + \|b\|)\,[ta_1 + (1 - t)b_1],$$

где

$$t = \frac{\|a\|}{\|a\| + \|b\|}, \quad a_1 = \frac{a}{\|a\|}, \quad b_1 = \frac{b}{\|b\|}.$$

силу аксиомы 2

$$\|a_1\| = \frac{1}{\|a\|}\|a\| = 1, \quad \|b_1\| = \frac{1}{\|b\|}\|b\| = 1,$$

$$\|a + b\| = (\|a\| + \|b\|)\,\|ta_1 + (1 - t)b_1\|. \tag{1}$$

В силу предположения сфера $\|x\| \leqslant 1$ выпукла. Поэтому вместе с a_1 и b_1 она должна заключать также и $ta_1 + (1-t) b_1$:

$$\| ta + (1-t) b_1 \| \leqslant 1. \tag{2}$$

Из (1) и (2) следует, что

$$\|a+b\| \leqslant \|a\| + \|b\|,$$

т. е. выполнена аксиома треугольника.

Линейные оболочки. В пространстве B задано множество X. Совокупность всевозможных линейных комбинаций

$$\sum_{i=1}^{n} \lambda_i x_i$$

элементов $x_i \subset X$ образует так называемую *линейную оболочку* над X.

Линейная оболочка над X есть наименьшее линейное многообразие, заключающее X.

Если замкнуть линейную оболочку, т. е. добавить к ней все элементы, предельные для ее последовательностей, то получится так называемая *замкнутая линейная оболочка* множества X.

Пример Пусть X есть последовательность степеней $1, t, t^2, \ldots, t^n, \ldots$ в C. Линейная оболочка множества X есть совокупность всех полиномов. В силу теоремы Вейерштрасса всякая $x(t) \subset C$ есть предел последовательности полиномов, следовательно, замкнутая линейная оболочка множества X совпадает со всем C.

Линейные функционалы и линейные операторы. Пусть E и E_1 — два пространства типа B. Оператор

$$y = s(x),$$

отображающий E на E_1, называется *линейным*, если

1) $$s(x_1 + x_2) = s(x_1) + s(x_2); \tag{3}$$

2) оператор s непрерывен:

$$s(x_n) \to s(x) \quad \text{при} \quad x_n \to x.$$

В частности, если E_1 есть числовая прямая, мы получаем *линейный функционал*.

Пример. Если $K(s, t)$ — непрерывная функция двух переменных, то оператор Фредгольма

$$y(t) = \int_0^1 K(s, t)\, x(s)\, ds$$

есть линейный оператор, отображающий C на C.

Теорема 1. *Если оператор $y = s(x)$ удовлетворяет условию* 1) *и непрерывен хотя бы в одной точке $x_0 \subset E$, то он есть линейный оператор.*

Нам нужно доказать, что $s(x)$ непрерывен во всякой точке $x \subset E$. Пусть $x_n \to x$; тогда

$$x_n + x_0 - x \to x_0.$$

Вследствие непрерывности s в точке x_0 имеем

$$s(x_n + x_0 - x) \to s(x_0). \tag{4}$$

Но

$$s(x_n + x_0 - x) = s(x_n) + s(x_0) - s(x). \tag{5}$$

Из (4) и (5) следует, что

$$s(x_n) \to s(x).$$

Теорема 2. *Линейный оператор $s(x)$ удовлетворяет условию однородности:*

$$s(\lambda x) = \lambda s(x). \tag{6}$$

В самом деле, для целого n имеем

$$s(nx) = s(\underbrace{x + x + \ldots + x}_{n \text{ раз}}) = \underbrace{s(x) + s(x) + \ldots + s(x)}_{n \text{ раз}} = ns(x). \tag{7}$$

Отсюда обозначая $nx = y$, имеем

$$s\left(\frac{1}{n}y\right) = \frac{1}{n}s(y). \tag{7'}$$

Комбинируя (7) и (7′), получаем

$$s\left(\frac{m}{n}x\right) = \frac{m}{n}s(x).$$

Итак, равенство (6) доказано для рациональных $\lambda = \dfrac{m}{n}$.

В силу непрерывности оператора s оно верно и для иррациональных λ.

Теорема 3. *Если оператор $y = s(x)$ линеен, то существует постоянная m такая, что*

$$\|s(x)\| \leqslant m\|x\|. \tag{8}$$

Допустим обратное, т. е. что отношение $\dfrac{\|s(x)\|}{\|x\|}$ неограниченно. Тогда для любого целого $n > 0$ найдется элемент $x_n \subset E$, для которого

$$\|s(x_n)\| > n\|x_n\|.$$

Обозначив: $x'_n = \dfrac{1}{n\|x_n\|}x_n$, получаем

$$\|x'_n\| = \frac{1}{n\|x_n\|}\|x_n\| = \frac{1}{n},$$

откуда

$$x'_n \to \theta_E.$$

Вследствие непрерывности s мы должны были бы иметь

$$s(x'_n) \to s(\theta_E) = \theta_{E_1}.$$

Но, с другой стороны, в силу теоремы 2

$$\|s(x'_n)\| = \frac{1}{n\|x_n\|}\|s(x_n)\| \geqslant 1.$$

Полученное противоречие доказывает теорему.

Замечание. Обратно, если оператор удовлетворяет условию аддитивности (3) и неравенству (8), то он линеен. В самом деле, из неравенства (8) следует, что при $x \to \theta_E$ $s(x) \to \theta_{E_1}$, т. е. в θ_E оператор $s(x)$ непрерывен. Следовательно, по теореме 1, он непрерывен всюду на E и, значит, линеен.

Норма линейного оператора. *Нормой линейного оператора* мы будем называть наименьшее из чисел m, удовлетворяющих неравенству (8). Обозначим его через $\|s\|$.

Имеем, с одной стороны,

$$\|s(x)\| \leqslant \|s\| \, \|x\|, \tag{9}$$

с другой стороны, при любом $\varepsilon > 0$ найдется элемент $x \subset E$, для которого

$$\|s(x)\| > (\|s\| - \varepsilon) \|x\|. \tag{10}$$

Теорема 4. *$\|s\|$ равна верхней грани $\|s(x)\|$ на сфере $\|x\| \leqslant 1$*:

$$\|s\| = \sup_{\|x\| \leqslant 1} \|s(x)\|.$$

В самом деле, при $\|x\| \leqslant 1$ имеем [в силу (9)]

$$\|s(x)\| \leqslant \|s\| \, \|x\| \leqslant \|s\|. \tag{11}$$

С другой стороны, при любом $\varepsilon > 0$ существуют элементы x, для которых имеет место неравенство (10). Обозначим $x' = \frac{x}{\|x\|}$; очевидно, $\|x'\| = 1$. Имеем в силу (10)

$$\|s(x')\| \geqslant (\|s\| - \varepsilon) \|x'\| = \|s\| - \varepsilon,$$

что вместе с (11) доказывает теорему.

В частности, для линейного функционала $f(x)$ неравенство (9) принимает вид

$$|f(x)| \leqslant \|f\| \, \|x\|,$$

где

$$\|f\| = \sup_{\|x\| \leqslant 1} |f(x)|.$$

Геометрическая интерпретация. Назовем *плоскостью* в пространстве E совокупность его точек x, удовлетворяющих уравнению

$$f(x) = c,$$

где f — линейный функционал.

Плоскости

$$f(x) = c, \quad f(x) = c_1$$

назовем *параллельными*. Плоскость $f(x) = c$ делит пространство на две части: совокупность точек x, в которых $f(x) > c$, и совокупность точек x, в которых $f(x) \leqslant c$. Мы условно назовем первое из этих полупространств лежащим *вправо*, второе — лежащим *влево* от плоскости $f(x) = c$. Плоскость $f(x) = \|f\|$ обладает тем свойством, что вся сфера $\|x\| \leqslant 1$ лежит целиком влево от этой плоскости [ибо для точек сферы $\|x\| \leqslant 1$ имеем $f(x) \leqslant \|f\|$]. С другой стороны, любая из параллельных плоскостей $f(x) = \|f\| - \varepsilon$, при $\varepsilon > 0$, этим свойством уже не обладает. По аналогии с теорией выпуклых n-мерных тел мы назовем плоскость $f = \|f\|$ *опорной* к сфере $\|x\| \leqslant 1$.

Пространство линейных операторов. Совокупность линейных операторов s, отображающих E на E_1, сама образует линейную систему Y: от сложения линейных операторов и умножения их на скаляр снова получаем линейный оператор. Нулевым элементом θ_Y можно считать оператор, отображающий E в нуль-элемент θ_{E_1}. В качестве нормы s можно принять

$$\|s\| = \sup_{\|x\|\leqslant 1} \|s(x)\|.$$

При этом аксиомы нормы удовлетворяются. Мы остановимся лишь на аксиоме треугольника.

Пусть s_1 и s_2 — операторы, отображающие E на E_1. Имеем

$$\|s_1\| = \sup_{\|x\|\leqslant 1} \|s(x)\|,$$

$$\|s_2\| = \sup_{\|x\|\leqslant 1} \|s_2(x)\|,$$

$$\|s_1 + s_2\| = \sup_{\|x\|\leqslant 1} \|s_1(x) + s_2(x)\|.$$

Так как для любого x

$$\|s_1(x) + s_2(x)\| \leqslant \|s_1(x)\| + \|s_2(x)\|, \tag{12}$$

то, переходя в неравенстве (12) к верхним границам по сфере $\|x\| \leqslant 1$, получаем

$$\|s_1 + s_2\| \leqslant \|s_1\| + \|s_2\|.$$

Особенно важную роль играет пространство линейных функционалов $f(x)$, определенных в E, где норма

$$\|f\| = \sup_{\|x\|\leqslant 1} |f(x)|.$$

Это пространство называется *сопряженным к* E и обозначается $\overline{E}$.

§ 6. Линейные функционалы и сопряженные пространства.

Теорема Гана. (Hahn). Рассмотрим в пространстве E линейное многообразие E'. Пусть в E' задан линейный функционал $f(x)$. Если в E задан линейный функционал $F(x)$, совпадающий в точках E' с $f(x)$:

$$F(x) = f(x) \quad \text{для} \quad x \subset E',$$

то F называется *распространением* f на все пространство E.

Теорема 1 (Гана). *Если в линейном подпространстве E' пространства E задан линейный функционал f, то его можно распространить на все пространство E без увеличения нормы f, т. е. можно построить на E линейный функционал F, совпадающий в E' с f, такой, что*

$$\|f\|_{E'} = \|F\|_E$$

(норма f по отношению к E' равна норме F по отношению к E).

Доказательство. Пусть E' — произвольное линейное многообразие в E, а x_0 — произвольный элемент в E. Обозначим через $(E' + x_0)$ совокупность всех элементов E вида

$$x + tx_0, \tag{1}$$

где $x \subset E, -\infty < t < \infty$. Очевидно, $(E' + x_0)$ есть тоже линейное многообразие в E (совпадающее с E', если $x_0 \subset E'$).

Если x_0 не входит в E', то каждый элемент $(E' + x_0)$ может быть единственным образом представлен в виде (1).

В самом деле, если

$$x + tx_0 = x_1 + t_1 x_0, \quad t \neq t_1, \quad x, x_1 \subset E',$$

то

$$x_0 = \frac{x - x_1}{t_1 - t}$$

и, следовательно, $x_0 \subset E'$ вопреки предположению. Поэтому $t_1 = t$ и $x = x_1$.

Мы покажем, что если линейный функционал $f(x)$ определен в E', то его можно распространить без увеличения нормы на $(E' + x_0)$. Очевидно, достаточно рассмотреть случай, когда x_0 не входит в E'.

Обозначим

$$\|f\|_{E_1} = M. \tag{2}$$

Имеем для любых $x', x'' \subset E'$

$$\begin{aligned} f(x') - f(x'') = f(x' - x'') \leqslant M \| x' - x'' \| = \\ = M \| (x' + x_0) - (x'' + x_0) \| \leqslant M \| x' + x_0 \| + M \| x'' + x_0 \| \end{aligned}$$

(последнее неравенство получается из аксиомы треугольника). Отсюда следует:

$$f(x') - M \| x' + x_0 \| \leqslant f(x'') + M \| x'' + x_0 \|.$$

Так как это неравенство имеет место для любой пары x', x'' из E', то

$$\sup \{ f(x') - M \| x' + x_0 \| \} \leqslant \inf \{ f(x'') + M \| x'' + x_0 \| \}. \tag{3}$$

Обозначим через c любое число, заключенное между правой и левой частями неравенства (3).

Такое число существует; в силу его определения для любых x' и x'' из E' имеем

$$f(x') - M \| x' + x_0 \| \leqslant c \leqslant f(x'') + M \| x'' + x_0 \|. \tag{4}$$

Определим теперь линейный функционал $F(x)$ в $(E' + x_0)$ следующим образом: если

$$x = x_1 + tx_0 \quad (x_1 \subset E'),$$

то

$$F(x) = f(x_1) - tc.$$

Очевидно, $F(x) = f(x)$ при $x \subset E'$, т. е. $F(x)$ есть распространение $f(x)$ на $(E' + x_0)$; остается доказать, что

$$\| F \|_{(E' + x_0)} = \| f \|_{E'}.$$

Пусть сначала $t > 0$. Для

$$x = x_1 + tx_0 = t \left(\frac{x_1}{t} + x_0 \right), \quad x_1 \subset E'$$

имеем

$$F(x) = f(x_1) - tc = t \left\{ f\left(\frac{x_1}{t} \right) - c \right\}.$$

Применяя к c первое из неравенств (4), полагая $x' = \frac{x_1}{t}$, получим

$$F(x) \leqslant t\left\{f\left(\frac{x_1}{t}\right) + \left[M\left\|x_0 + \frac{x_1}{t}\right\| - f\left(\frac{x_1}{t}\right)\right]\right\} =$$
$$= Mt\left\|x_0 + \frac{x_1}{t}\right\| = M\|x_1 + tx_0\| = M\|x\|. \tag{5}$$

Аналогично, при отрицательном t мы пришли бы к тому же неравенству

$$F(x) \leqslant M\|x\|,$$

используя второе из неравенств (4) и полагая $x'' = \frac{x_1}{t}$.

Итак, неравенство (5) справедливо для всех $x \subset (E' + x_0)$. Меняя знак при x (отчего $\|x\|$ не изменится), получим

$$F(x) \geqslant -M\|x\|,$$

тоже для всех x. В соединении с (5) и (2) это дает

$$\|F\|_{(E'+x_0)} \leqslant M = \|f\|_{E'}.$$

Но так как, очевидно, при распространении линейного функционала его норма (т. е. его максимум на сфере радиуса 1) не может уменьшиться, то

$$\|F\|_{(E'+x_0)} = \|f\|_{E'},$$

что и требовалось доказать.

Теперь можно доказать и всю теорему Гана.

Рассмотрим сначала случай, когда E сепарабельно. Пусть $(x_1, x_2, \ldots, x_n, \ldots)$— счетная всюду плотная сеть на E. Как мы видели, f можно последовательно распространить на пространства $(E' + x_1)$, $(E' + x_1 + x_2)$, $(E' + x_1 + x_2 + x_3)$ и т. д., не увеличивая его нормы. После счетного множества распространений мы распространим f, не увеличивая его нормы, на подпространство $(E' + x_1 + x_2 + \ldots)$, всюду плотное на E. А затем предельным переходом мы распространим f, не увеличивая его нормы, на все пространство E.

Для несепарабельного случая нам придется вместо счетной последовательности $\{x_n\}$, плотной на E, рассмотреть такую же трансфинитную последовательность $\{x_\alpha\}$ (высшей мощности), и тем самым вместо счетной последовательности распространений f рассмотреть трансфинитную последовательность высшей же мощности. К сожалению, мы не обладаем другим доказательством теоремы Гана, исключающим подобные трансфинитные рассмотрения.

Как следствие этой теоремы получаем:

Теорема 2. *Для любого элемента x_0 существует такой линейный функционал f_0, что*

$$f_0(x_0) = \|f_0\|\,\|x_0\|.$$

Можно полагать норму $\|f_0\|$ равной единице, и тогда

$$f_0(x_0) = \|x_0\|.$$

[Заметим, что для любого линейного функционала f имеет всегда место неравенство

$$f(x_0) \leqslant \|f\|\,\|x_0\|.$$

Из (10) и (11) имеем

$$|f_m^{(m)}(x)-f_p^{(p)}(x)|\leqslant|f_m^{(m)}(x)-f_m^{(m)}(x_n)|+|f_m^{(m)}(x_n)-f_p^{(p)}(x_n)|+$$
$$+|f_p^{(p)}(x_n)-f_p^{(p)}(x)|\leqslant\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.$$

Отсюда следует сходимость $f_k^{(k)}(x)$ при $k\to\infty$ для любого x.

Итак, мы выделили из первоначальной последовательности f_n подпоследовательность $f_k^{(k)}$, сходящуюся для любого $x \subset E$, т. е. слабо сходящуюся. Теорема доказана.

Рассматривая, например, слабую сходимость в C_M, получим теорему, которую Радон доказывает в приводимой ниже статье, на стр. 206—207.

§ 9. Универсальность пространства C.

П. С. Урысон доказал существование „универсального" метрического сепарабельного пространства, т. е. такого, которое содержит части, изометричные любому сепарабельному метрическому пространству. Впоследствии Банах и Мазур доказали, что C является таким универсальным пространством.

Доказательство этой теоремы связано со свойством слабой компактности сопряженных пространств.

Обозначим через U совокупность всех функционалов $f \subset \overline{E}$, для которых $\| f \| \leqslant 1$, причем сходимость в U совпадает со слабой сходимостью функционалов.

Пусть E сепарабельно, и $x_1, x_2, \ldots, x_n, \ldots$ — плотная сеть на единичной сфере $\| x \| \leqslant 1$. Для слабой сходимости $f^{(k)} \overset{\text{сл}}{\to} f$, достаточно, чтобы при $k\to\infty$

$$f^{(k)}(x_n)\to f(x_n), \quad n=1,\ 2,\ 3,\ldots$$

Таким образом U можно рассматривать как часть пространства N (см. § 1, пример 6) и притом компактную часть. Каждому элементу $f \subset U$ отвечает элемент из N с координатами $f(x_i)$, $i=1,\ 2,\ldots,\ n,\ldots$

Заметим, что N, а следовательно и U, сепарабельно. Существует следующая теорема.

Теорема 1 Гаусдорфа. *Всякое компактное сепарабельное множество есть непрерывный образ замкнутого множества P отрезка* [0, 1].

Мы проведем доказательство этой теоремы для нашего частного случая.

U есть замкнутое множество точек $x(x_1, x_2, \ldots, x_n, \ldots)$ из N, где все координаты $|x_i|\leqslant 1$. Пусть мы имеем точку $x \subset N$ с координатами $x_i, |x_i|\leqslant 1$. Представим числа x_i по двоичной системе:

$$\left.\begin{array}{l} x_1=0,\ \alpha_{11}\,\alpha_{12}\,\alpha_{13}\,\alpha_{14}\ldots \\ x_2=0,\ \alpha_{21}\,\alpha_{22}\,\alpha_{23}\,\alpha_{24}\ldots \\ x_3=0,\ \alpha_{31}\,\alpha_{32}\,\alpha_{33}\,\alpha_{34}\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\} \tag{1}$$

(α_{ik} — цифры 0 или 1), и отнесем нашей точке x число y, которое имеет по троичной системе разложение

$$y=0,\ \alpha_{11}\,\alpha_{12}\,\alpha_{21}\,\alpha_{13}\,\alpha_{22}\,\alpha_{31}\,\alpha_{14}\,\alpha_{23}\,\alpha_{32}\,\alpha_{41}\ldots \tag{2}$$

(все цифры в троичном представлении y равны нулю или единице).

Обратно, каждому числу y, представимому по троичной системе в виде (2) (т. е. обладающему лишь цифрами 0 и 1), отвечает точка $x \subset N$, составленная по схеме (1).

При таком отнесении совокупности $U \subset N$ точек $x(x_i)$ отвечает некоторая совокупность Y чисел y на отрезке $[0, 1]$. При этом одной точке $x(x_1, x_2, \ldots, x_n, \ldots)$ может отвечать целая система чисел y (если координата x_i точки x двоично рациональна, то она допускает два представления по двоичной системе). Наоборот, число $y \subset Y$ допускает единственное представление по троичной системе с помощью цифр 0 и 1 (если y троично рационально, то мы должны ограничиться только тем его представлением, в котором все цифры, начиная с некоторой, равны нулю).

Таким образом каждой $y \subset Y$ отвечает точка $x = \varphi(y) \subset U$, причем соответствие $x = \varphi(y)$ *однозначно*. Докажем, что это соответствие непрерывно.

Пусть $y_n \to y$, где y_n и y представимы по троичной системе в виде (2) с помощью цифр 0 и 1. Число первых цифр, совпадающих в троичных разложениях чисел y_n и y, неограниченно растет при $n \to \infty$. Сопоставим соответственные точки из N:

$$x^{(n)} = \varphi(y^{(n)}) \quad \text{и} \quad x = \varphi(y),$$

где $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}, \ldots)$, $x = (x_1, x_2, \ldots, x_k, \ldots)$.

В силу построения $x^{(n)}$ и x число первых цифр, совпадающих у координат $x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}, \ldots$ точек $x^{(n)}$ и соответствующих координат $x_1, x_2, \ldots, x_k, \ldots$ точки x неограниченно растет, а значит, $x_1^{(n)} \to x_1$, $x_2^{(n)} \to x_2, \ldots$ при $n \to \infty$; по определению сходимости в N, $x^{(n)} \to x$.

Итак, отображение $x = \varphi(y)$ непрерывно.

Докажем, что множество Y замкнуто на отрезке $[0, 1]$. Пусть последовательность $y_n \subset Y$ сходится к числу y. Числа y_n, равно как и предельное число y, допускают троичное представление с помощью цифр 0, 1.

В силу непрерывности оператора φ имеем $x^{(n)} \to x$, где $x^{(n)} = \varphi(y_n)$, $x = \varphi(y)$. Так как $y_n \subset Y$, то, по определению, $x^{(n)} \subset U$. Но U замкнуто, следовательно, $x \subset U$. Поэтому и $y \subset Y$.

Итак, существует однозначное непрерывное отображение замкнутого множества $Y \subset [0, 1]$ на $U \subset N$. Приступим к доказательству теоремы Банаха. Докажем предварительно следующую теорему.

Теорема 2 (Фреше). *Всякое метрическое сепарабельное пространство M изометрично части некоторого линейного сепарабельного пространства.*

Доказательство. Пусть $x_0, x_1, x_2, \ldots, x_n, \ldots$ образуют счетную всюду плотную сеть на M. Отнесем каждому элементу $x \subset M$ точку $y(y_i)$ пространства m (см. § 1, пример 5), где

$$y_i = \rho(x, x_i) - \rho(x_0, x_i), \quad i = 0, 1, 2, \ldots$$

Заметим, что в силу аксиомы треугольника

$$|y_i| = |\rho(x, x_i) - \rho(x_0, x_i)| \leqslant \rho(x_0, x),$$

таким образом числа y_i ограничены, и, следовательно, $y(y_i)$ действительно можно рассматривать как точку из m.

Мы доказываем, что для каждого x_0 найдется функционал $f=f_0$, обращающий это неравенство в равенство.]

Обозначим через E' прямую, составленную из элементов tx_0, $-\infty<t<\infty$. Определим на E' линейный функционал

$$f(tx_0)=t\,\|x_0\|,$$

так что в частности

$$f(x_0)=\|x_0\|.$$

Очевидно, $\|f\|_{E'}=1$. В силу предыдущей теоремы можно, не увеличивая нормы f распространить f на все пространство.

Геометрически теорема заключается в следующем: *через всякую точку x_0 расположенную на границе сферы $\|x\|\leqslant 1$ (т. е. такую, что $\|x_0\|=1$), можно провести опорную плоскость к этой сфере.*

В самом деле, уравнение опорной плоскости к единичной сфере должно иметь вид

$$f(x)=\|f\|.$$

Но для точки x_0 можно построить функционал f_0 с нормой 1, для которого

$$f_0(x_0)=\|x_0\|=1. \tag{6}$$

Опорная плоскость

$$f_0(x)=\|f_0\|=1$$

проходит в силу (6) через точку x_0.

Предел последовательности линейных функционалов. В дальнейшем нам понадобится еще следующая теорема.

Теорема 3. *Если последовательность $f_n(x)$ линейных функционалов на E сходится для любого $x\subset E$, то*

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

тоже есть линейный функционал.

Доказательство. Имеем

$$f(x+y)=\lim_{n\to\infty}f_n(x+y)=\lim_{n\to\infty}[f_n(x)+f_n(y)]=f(x)+f(y).$$

Итак, $f(x)$ удовлетворяет условию аддитивности. Остается доказать, что $f(x)$ — непрерывный функционал. В силу теоремы 1, § 5, достаточно доказать, что $f(x)$ непрерывен хотя бы в одной точке.

Для функций вещественного переменного, являющихся пределами непрерывных функций, т. е. для так называемых функций первого класса, Бэр доказал, что они обладают одной по крайней мере точкой непрерывности[1]). Теорема Бэра (и ее доказательство) распространяются и на функционалы, определенные на любом метрическом пространстве. В силу этого функционал $f(x)$, являющийся пределом *непрерывных* функционалов $f_n(x)$, должен обладать одной по крайней мере точкой непрерывности, чем теорема и доказана.

---

[1]) См. Р. Бэр, Разрывные функции.

Заметим, что теорема Бэра, а, следовательно, и доказанная только что теорема, распространяется и на операторы.

Общий вид линейных функционалов. 1. В n-мерных евклидовых пространствах E_n линейный функционал $f(x)$ имеет вид

$$f(x)=f(x_1, x_2, \ldots, x_n)=\sum_{i=1}^{n} f_i x_i.$$

В силу неравенства Шварца

$$|f(x)|\leqslant\sqrt{\sum_{i=1}^{n} f_i^2 \sum_{i=1}^{n} x_i^2}\leqslant\sqrt{\sum_{i=1}^{n} f_i^2}\,\|x\|,$$

причем равенство достигается (для x_i, пропорциональных к f_i).

Отсюда

$$\|f\|=\sqrt{\sum_{i=1}^{n} f_i^2}. \tag{7}$$

Каждый функционал $f\subset\overline{E}_n$ определяется n числами ($f_1, f_2, \ldots, f_n$). Таким образом пространство $\overline{E}_n$ есть n-мерное пространство, причем в силу формулы (7) оно имеет также евклидову метрику. Отсюда: *пространство $\overline{E}_n$, сопряженное к евклидову, есть то же самое евклидово пространство.* Иными словами евклидово пространство является *самосопряженным.*

2. Рассмотрим линейный функционал $f(x)$ в гильбертовом пространстве l_2.

Обозначим через X_n совокупность точек $x\subset l_2$ вида

$$x=(x_1, x_2, \ldots, x_n, 0, 0, \ldots, 0, \ldots).$$

X_n есть евклидово n-мерное пространство, и на нем $f(x)$ имеет вид

$$f(x)=\sum_{i=1}^{n} f_i x_i.$$

Поэтому

$$|f(x)|\leqslant\sqrt{\sum_{i=1}^{n} f_i^2}\sqrt{\sum_{i=1}^{n} x_i^2}=\sqrt{\sum_{i=1}^{n} f_i^2}\,\|x\|,$$

причем равенство достигается.

Отсюда в силу определения $\|f\|$ имеем

$$\sqrt{\sum_{i=1}^{n} f_i^2}\leqslant\|f\|.$$

Будем переходить к пределу; в силу ограниченности сумм $\sum\limits_{i=1}^{n} f_i^2$ ряд $\sum\limits_{i=1}^{n} f_i^2$ сходится. Каждый элемент $x(x_1, x_2, \ldots, x_n, \ldots)$ есть предел элементов $x^{(n)}(x_1, x_2, \ldots, x_n, 0, \ldots)\subset X_n$. Отсюда

$$f(x)=\lim_{n\to\infty} f(x^{(n)})=\lim_{n\to\infty}\left(\sum_{i=}^{n} f_i x_i\right).$$

Так как ряды $\sum\limits_{i=1}^{\infty} x_i^2$ и $\sum\limits_{i=1}^{\infty} f_i^2$ сходятся, то в силу неравенства Шварца ряд

$$\sum_{i\ 1}^{\infty} f_i x_i$$

тоже сходится. Имеем

$$f(x)=\sum_{i=1}^{\infty} f_i x_i.$$

Применяя неравенство Шварца, получаем

$$|f(x)| \leqslant \sqrt{\sum_{i=1}^{\infty} f_i^2 \sum_{i=1}^{\infty} x_i^2},$$

причем равенство достигается, если x_i пропорциональны f_i. Отсюда

$$\|f\|=\sqrt{\sum_{i=1}^{\infty} f_i^2}.$$

Каждый функционал $f(x)$ определяется последовательностью чисел

$$(f_1, f_2, \ldots, f_n, \ldots)$$

со сходящейся суммой квадратов, причем

$$\|f\|=\sqrt{\sum_{i=}^{\infty} f_i^2}.$$

Таким образом функционал f есть сам элемент гильбертова пространства. *Пространство, сопряженное к гильбертову пространству, совпадает с ним самим*; гильбертово пространство также является самосопряженным.

3. Аналогичное предложение можно доказать и для функционального пространства L_2. Можно показать, что любой линейный функционал $f(y)$ на L_2 представим в виде

$$f(y)=\int_0^1 y(t)\varphi(t)\,dt,$$

где $\varphi(t)$ есть функция с интегрируемым квадратом, причем $\|f\|=\sqrt{\int_0^1 \varphi^2\,dt}$.

Таким образом

$$\overline{L}_2=L_2,$$

что, впрочем, вытекает из самосопряженности l_2, так как L_2 ему изометрично [1].

4. В пространстве $l_p^{(n)}$, $p>1$, функционал $f(x)$ имеет вид

$$f(x)=\sum_{i=1}^{n} f_i x_i.$$

Применим неравенство Гельдера:

$$\left|\sum_{i=}^{n} f_i x_i\right| \leqslant \left\{\sum_{i=1}^{n}|f_i|^q\right\}^{\frac{1}{q}}\left\{\sum_{i=1}^{n}|x_i|^p\right\}^{\frac{1}{p}}=\left\{\sum_{i=1}^{n}|f_i|^q\right\}^{\frac{1}{q}}\|x\|,$$

где $\frac{1}{p}+\frac{1}{q}=1$ или $q=\frac{p}{p-1}$ (причем равенство достигается).

[1]) До сих пор остается открытым вопрос, существуют ли самосопряженные сепарабельные пространства, неизометричные гильбертову.

Отсюда

$$\|f\| = \left\{ \sum_{i=1}^{n} |f_i|^q \right\}^{\frac{1}{q}}.$$

Итак, функционал $f \subset \overline{l_p^{(n)}}$ есть элемент n-мерного пространства $l_q^{(n)}$:

$$\overline{l_p^{(n)}} = l_q^{(n)}.$$

Аналогично, путем предельного перехода, доказывается, что

$$\begin{array}{l} \overline{l_p} = l_q, \\ \overline{L_p} = L_q \end{array} \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right).$$

В частности, при $p = 2$, $q = 2$, получаем предыдущий результат.

5. В пространстве C непрерывных функций $y(t)$ (определенных на [0, 1]) линейный функционал выражается в виде интеграла Стилтьеса:

$$f(y) = \int_0^1 y(t)\, dg(t),$$

где $g(t)$ — функция с ограниченным изменением и

$$\|f\| = \operatorname{var} g(t),$$

причем через $\operatorname{var} g(t)$ обозначено полное изменение $g(t)$ на [0, 1].

Это составляет содержание теоремы Рисса. Мы ее здесь не доказываем, потому что в следующем параграфе докажем более общую теорему.

Полнота. $\overline{E}$.

Теорема 4. *Всякое сопряженное пространство $\overline{E}$ есть полное пространство.*

Пусть $f_1, f_2, \ldots, f_n, \ldots$ — фундаментальная последовательность в $\overline{E}$. Для любого $m > 0$

$$\|f_{n+m} - f_n\| \leqslant \varepsilon_n, \tag{8}$$

где $\varepsilon_n \to 0$ при $n \to \infty$. Из (8) следует

$$|f_{n+m}(x) - f_n(x)| \leqslant \varepsilon_n \|x\|, \tag{9}$$

где x — любой элемент из E.

Последовательность $f_n(x)$ для любого фиксированного x есть фундаментальная числовая последовательность. Следовательно, $f_n(x)$ сходится для любого x.

Обозначим

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

Имеем

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

Далее, так как нормы f_n ограничены, $\|f_n\| < K$, то

$$|f_n(x)| \leqslant K\|x\|;$$

и переходя к пределу по $n \to \infty$, получаем

$$|f(x)| \leqslant K\|x\|.$$

В силу замечания к теореме 3, § 5, $f(x)$ есть линейный функционал.

Переходя в неравенстве (9) к пределу по $m \to \infty$, получим

$$|f(x) - f_n(x)| \leqslant \varepsilon_n \|x\|, \quad \text{т. е.} \quad \|f_n - f\| \leqslant \varepsilon_n,$$

откуда

$$f_n \to f.$$

Применение. Поскольку пространства L_p $(p > 1)$ суть сопряженные к L_q $\left(\frac{1}{q} + \frac{1}{p} = 1\right)$, то мы получаем новое доказательство полноты L_p.

Внутренние произведения. Рассмотрим выражение $f(x)$, где $f \subset \overline{E}$, $x \subset E$. Если рассматривать x как переменную величину, а f фиксировать, то $f(x)$ будет линейным функционалом в E. Если же x фиксировать и менять $f \subset \overline{E}$, то выражение $f(x)$ обратится в функционал, определенный на $\overline{E}$.

Ввиду такой симметрии мы будем писать выражение $f(x)$ без скобок:

$$fx = f(x). \tag{10}$$

fx есть *билинейный функционал*, он зависит линейно и от $f \subset \overline{E}$ и от $x \subset E$. Этот билинейный функционал можно было бы по аналогии с векторной алгеброй называть *внутренним произведением* f и x. Для самосопряженных пространств (евклидово, гильбертово) $\overline{E} = E$ и оба аргумента в (10) принадлежат одному пространству.

Если рассматривать (10) при постоянном f, то получим функционал в E, норма которого равна $\|f\|$.

Фиксировав x, получаем функционал в $\overline{E}$, который по аналогии обозначим $x(f)$

$$x(f) = fx. \tag{11}$$

Определим норму этого функционала. Так как, с одной стороны, имеем

$$|x(f)| = |fx| \leqslant \|f\| \|x\|,$$

а, с другой стороны, для данного $x \subset E$ существует (теорема 2) такое $f \subset \overline{E}$, что

$$|x(f)| = |fx| = \|f\| \|x\|,$$

то норма функционала $x(f)$ в $\overline{E}$ равна $\|x\|$.

Рассматривая совокупность всех линейных функционалов $L(f)$ в $\overline{E}$ как новое линейное пространство $\overline{\overline{E}}$, получим: $\overline{\overline{E}}$ заключает в себе E, ибо каждый x порождает линейный функционал $x(f)$ на $\overline{E}$ (т. е. элемент $\overline{\overline{E}}$), причем норма этого функционала (норма в $\overline{\overline{E}}$) совпадает с $\|x\|$, нормой x в E[1]). Если бы всякий линейный функционал $L(f)$ в $\overline{E}$ имел вид (11), то $\overline{\overline{E}}$ совпадало бы с E. Такие пространства называются *регулярными*.

---

[1]) Говоря, что $\overline{\overline{E}}$ заключает E, мы понимаем это именно в том смысле, что E изометрично (см. выше, стр. 86) части пространства $\overline{\overline{E}}$.

Пример. 1. *Все конечномерные линейные пространства регулярны.*

В самом деле, если E есть совокупность n-мерных векторов $x(x_1, x_2, \ldots, x_n)$, то линейный функционал $l(x)$ имеет вид

$$l(x) = \sum_{i=1}^{n} l_i x_i.$$

$\overline{E}$ состоит из n-мерных векторов $l(l_1, l_2, \ldots, l_n)$. Пусть мы имеем линейный функционал $F(l)$ в $\overline{E}$. Он имеет вид

$$F(l) = \sum_{i=1}^{n} x_i l_i = xl,$$

где $x = (x_1, x_2, \ldots, x_n)$.

Пример. 2. *Пространства l_p и L_p регулярны.*

Пусть $\frac{1}{p} + \frac{1}{q} = 1$. Рассмотрим билинейную форму

$$J(y, z) = \sum_{i=}^{\infty} y_i z_i,$$

где $y(y_i) \subset l_p$, $z(z_i) \subset l_q$. Линейные функционалы как в l_p, так и в $\overline{l_p} = l_q$ выражаются формой $J(y, z)$ при фиксированных z или соответственно y. Отсюда следует регулярность l_p.

Аналогично доказывается регулярность L_p.

Пример. 3. *Пространство C нерегулярно.*

В самом деле, если бы $\overline{\overline{C}} = C$, то всякий линейный функционал $F(g)$ на $\overline{C}$ мог бы быть представлен в виде

$$F(g) = \int_0^1 x(t)\, dg(t),$$

где $x(t)$ — некоторая функция из C, для каждого функционала F своя, причем это равенство было бы справедливо, при заданных F и x, для всех g (с ограниченным изменением).

Рассмотрим линейный в $\overline{C}$ функционал $F(g)$, равный скачку функции $g(t)$ в какой-нибудь точке t_0: $F(g) = g(t_0 + 0) - g(t_0 - 0)$. $F(g)$ есть линейный функционал, не тождественно равный нулю, и если бы $C = \overline{\overline{C}}$, то он представлялся бы в виде

$$\int_0 x(t)\, dg(t),$$

где $x(t) \not\equiv 0$, $x(t) \subset C$. С другой стороны, для всех непрерывных $g(t) \subset C$ в частности для

$$g_0(t) = \int_0^t x(t)\, dt,$$

$$F(g_0) = 0.$$

Но

$$\int_0^1 x(t)\, dg_0(t) = \int_0^1 [x(t)]^2\, dt > 0$$

Полученное противоречие опровергает гипотезу о регулярности C.

А. И. Плеснер доказал, что или $\overline{\overline{E}} = E$ (пространство регулярно), или все пространства E, $\overline{E}$, $\overline{\overline{E}}$, $\overline{\overline{\overline{E}}}$, ... различны между собой (т. е. попарно неизометричны).

Сопряженный оператор. Мы сейчас будем рассматривать операторы, отображающие пространство E на все пространство E_1.

Пусть дан линейный оператор s, отображающий пространство E на E_1:

$$y = s(x), \quad x \subset E, \quad y \subset E_1 .$$

Каждый функционал $f(y) \subset \overline{E}_1$, определенный в E_1, переходит при этом в функционал $\varphi(x)$, определенный в E:

$$\varphi(x) = f[s(x)].$$

Будем писать

$$\varphi = \overline{s}(f).$$

Оператор $\overline{s}$ будем называть *сопряженным* к s.

В частности, каждому оператору s, отображающему E на самое себя, отвечает сопряженный оператор $\overline{s}$, отображающий $\overline{E}$ на самое себя.

Пример. Пусть E есть n-мерное пространство, E_1 — k-мерное пространство ($k \leqslant n$). Если $y(y_1, y_2, \ldots, y_k) \subset E_1$ и $x(x_1, x_2, \ldots, x_n) \subset E$, то оператор

$$y = s(x),$$

отображающий E на E_1, имеет вид

$$y_i = \sum_{j=1}^{n} s_{ij} x_j \quad (i = 1, 2, \ldots, k). \tag{12}$$

$\overline{E}_1$ есть совокупность $l(l_1, l_2, \ldots, l_k)$ линейных функционалов в E_1:

$$l(y) = \sum_{i=1}^{k} l_i y_i .$$

Имеем

$$l(y) = \sum_{i=1}^{k} l_i y_i = \sum_{i=1}^{k} l_i \left(\sum_{j=1}^{n} s_{ij} x_j \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{k} l_i s_{ij} \right) x_j = l'(x),$$

где l' есть функционал в E:

$$l'(x) = \sum_{j=1}^{n} l_j' x_j .$$

Оператор

$$l' = \overline{s}\, l$$

имеет вид

$$l_j' = \sum_{i=1}^{k} s_{ij} l_i \quad (j = 1, 2, \ldots, n),$$

или, заменяя друг другом индексы i и j,

$$l_i' = \sum_{j=1}^{k} s_{ji} l_j \quad (i = 1, 2, \ldots, n). \tag{13}$$

Матрица преобразования (13) сопряжена матрице преобразования (12). Итак, для конечномерного случая переход от оператора s к сопряженному оператору $\bar{s}$ приводит к переходу от данной матрицы преобразования к сопряженной.

Если мы имеем операторное уравнение

$$s(x) = y, \quad x \subset E, \quad y \subset E_1,$$

то сопряженное ему будет

$$\bar{s}(f) = \varphi, \quad f \subset \bar{E}_1, \quad \varphi \subset \bar{E}.$$

Связь между операторным уравнением и его сопряженным была впервые исследована (для $E = C$) в работе Радона, перевод которой печатается ниже (стр. 200—227); мы отсылаем читателя к этой работе.

См. также Banach, Théorie des opérations linéaires, 1932, гл. X, § 1.

§ 7. Теорема Рисса-Радона.

Задачей настоящего параграфа является нахождение общего вида линейного функционала в пространстве непрерывных функций. Будем обозначать через C_M линейное пространство, составленное из функций $f(A)$, определенных в точках пространства M и непрерывных на нем. При этом положим

$$\|f\| = \sup_{A \subset M} |f(A)|. \tag{1}$$

(В случае $M = [0, 1]$ мы получаем обычное пространство C.) В дальнейшем мы всюду под M будем понимать измеримое множество в n-мерном пространстве.

Определим еще пространство L_M, состоящее из всех ограниченных измеримых функций $f(A)$ на M, причем норма $\|f\|$ определяется также по формуле (1).

Абсолютно аддитивные функции множеств. Рассмотрим в метрическом пространстве M совокупность (E) множеств E, удовлетворяющую следующим условиям:

1) M входит в (E).

2) Сумма конечной системы множеств из (E) входит в (E).

3) Дополнение $M - E$ к множеству $E \subset (E)$ также входит в (E).

Будем в дальнейшем считать, что M совпадает с n-мерным пространством или каким-нибудь открытым или замкнутым множеством на нем, а совокупность (E) включает все области на M, а, значит, и все замкнутые множества на M.

Функция множества E, определенная на (E), относит каждому $E \subset (E)$ некоторое вещественное число $\varphi(E)$. Это число $\varphi(E)$ называется *аддитивной функцией множеств*, если для любого конечного разбиения множества E

$$E = E_1 + E_2 + \ldots + E_n,$$

где $E_i \subset (E)$ и все E_i попарно не пересекаются, будет

$$\varphi(E) = \varphi(E_1) + \varphi(E_2) + \ldots + \varphi(E_n).$$

Аддитивная функция $\varphi(E)$ называется *абсолютно аддитивной*, если для любой системы множеств $E_1, E_2, \ldots, E_n$ из (E), не имеющих попарно общих точек

$$\sum_{i=1}^{n} |\varphi(E_i)| \leqslant K, \tag{2}$$

где K — абсолютная постоянная[1].

---

[1] Более подробно об аддитивных функциях множеств см. В. И. Гливенко, Интеграл Стилтьеса, ОНТИ, 1936.

Пример. Если M имеет ограниченную меру, то лебегова мера $\operatorname{mes} E$ есть абсолютно аддитивная функция множеств.

Совокупнос ь абсолютно аддитивных функций множеств образует линейную систему; ее можно рассматривать как линейное пространство $\mathfrak{E}$, если в качестве нормы $\varphi(E) \subset \mathfrak{E}$ примем число $m(\varphi)$ — нижнюю границу чисел K в неравенстве (2): $m(\varphi) = \inf K$.

Рассмотрим, например, случай, когда M есть отрезок $[0, 1]$ числовой прямой. Обозначим через E_t отрезок $[0, t]$. Пусть задана на $[0, 1]$ абсолютно аддитивная функция множеств $\varphi(E)$ и пусть $g(t) = \varphi(E_t)$. Легко видеть, что $g(t)$ есть функция с ограниченным изменением, и что $m(\varphi)$ совпадает с полным изменением $g(t)$ на отрезке $[0, 1]$.

Сделаем в заключение несколько замечаний об абсолютно аддитивных функциях множеств.

Мы скажем: $\varphi(E)$ есть *положительно аддитивная функция множеств*, если сверх общих условий аддитивности она удовлетворяет еще условию положительности

$$\varphi(E) \geqslant 0.$$

Очевидно, положительно аддитивная функция есть и абсолютно аддитивная.

Теорема 1. *Всякая абсолютно аддитивная функция множеств $\varphi(E)$ есть разность двух положительно аддитивных функций.*

Определим:

$$\varphi_1(E) = \int\limits_E |d\varphi|$$

или, что то же самое,

$$\varphi_1(E) = \sup \sum_i |\varphi(E_i)|$$

для произвольных разбиений $E = E_1 + E_2 + \ldots + E_k$, где $E_i \subset E$, $i = 1, 2, \ldots, k$ и все E_i не имеют попарно общих частей. $\varphi_1(E)$ называется *полным колебанием* φ на E.

В силу абсолютной аддитивности $\varphi(E)$, $\varphi_1(E)$ определена для любого $E \subset (E)$; при этом, очевидно, $\varphi(E) \leqslant \varphi_1(E)$.

Но $\varphi_1(E)$ — положительно аддитивна. Поэтому и $\varphi_2(E) = \varphi_1(E) - \varphi(E)$ тоже есть положительно аддитивная функция множеств. Но тем самым мы нашли представление $\varphi(E)$ в виде разности двух положительно аддитивных функций множеств:

$$\varphi(E) = \varphi_1(E) - \varphi_2(E).$$

φ-измеримые множества. Напомним определение лебеговской меры множества E. Мы рассматриваем верхнюю границу мер заключенных в E замкнутых множеств. Далее, мы рассматриваем нижнюю границу заключающих E открытых множеств. Если оба эти числа совпадают, то их общее значение называется *мерой* E, а само E называется *измеримым по Лебегу* множеством.

Пусть для открытых и замкнутых множеств E' определена положительно аддитивная функция $\varphi(E')$. По аналогии с множествами, измеримыми по Лебегу, можно определить φ-измеримые множества E как те, для которых

$$\sup \varphi(E') = \inf \varphi(E''),$$

где E' и E'' суть соответственно замкнутые и открытые множества, причем E' заключены в E, а E'' заключают E; будем считать $\varphi(E)$ равным общему значению $\sup\varphi(E')$ и $\inf\varphi(E'')$.

Интеграль Стилтьеса-Радона. Пусть $\varphi(E)$ — абсолютно аддитивная функция множеств, $E \subset M$, и $F(A)$ — непрерывная функция точки на M. Обозначим через S разбиение M на n φ-измеримых частей без общих точек,

$$S: \qquad M = E_1 + E_2 + \ldots + E_n,$$

а через $\delta(S)$ — наибольший из диаметров множеств E_i, $i = 1, 2, \ldots, n$. Зададим систему разбиений $S_1, S_2, \ldots, S_k, \ldots$,

$$S_k: \qquad M = \sum_{i=1}^{n_k} E_i^{(k)},$$

для которых $\delta(S_k) \to 0$. Рассмотрим сумму

$$\sum_{i=1}^{n_k} F(A_i^{(k)})\,\varphi(E_i^k), \tag{3}$$

где $A_i^{(k)}$ — произвольная точка множества $E_i^{(k)}$.

Можно показать (подобно тому как мы это делаем в обычной теории интеграла Римана), что выражения (3) стремятся при $k \to \infty$ к пределу, который мы сбозначим через

$$\int_M F(A)\,d\varphi. \tag{4}$$

Выражение (4) не зависит ни от выбора точек $A_i^{(k)}$ на множествах $E_i^{(k)}$ в (3), ни от выбора разбиений S_k [лишь бы $\delta(S_k)$ стремилось к нулю].

Естественно положить

$$m(\varphi) = \int_M |d\varphi|. \tag{5}$$

Из определений (4) и (5) следует

$$\left|\int_M F(A)\,d\varphi\right| \leqslant \max|F(A)| \int_M |d\varphi|. \tag{6}$$

Линейные функционалы на C_M. Рассмотрим пространство C_M всех непрерывных функций $F_{(A)}$ на M, где $\|F\| = \max\limits_{A \subset M}|F(A)|$.

Теорема 2. (Рисса-Радона). *Всякий линейный функционал $l(F)$ на C_M имеет вид*

$$\int_M F(A)\,d\varphi, \tag{7}$$

где φ — абсолютно аддитивная функция множеств на M, причем

$$\|l\| = m(\varphi).$$

Таким образом определенное выше пространство абсолютно аддитивных функций φ есть пространство, сопряженное с C_M, $E = \overline{C}_M$[1]).

---

[1]) Здесь „равенство" пространств E и $\overline{C}_M$ понимается, конечно, как и всюду в аналогичных случаях, как их *изометричность*.

Доказательство. Обозначим через $\chi_E(A)$ (E — произвольное открытое или замкнутое множество, $E \subset M$) функцию, равную единице в точках A множества E и равную нулю вне E.

Распространим функционал $l(F)$, определенный в C_M, на пространство L_M. В силу теоремы Гана такое распространение можно произвести без увеличения нормы $\| l \|$.

Заметим, что все $\chi_E(A)$ входят в L_M, причем $\| \chi_E \| = 1$. Обозначим

$$\varphi(E) = l(\chi_E).$$

Лемма. $\varphi(E)$ есть абсолютно аддитивная функция множеств.

В самом деле, если E и E_1 не имеют общих точек, то по определению χ_E, χ_{E_1}, χ_{E+E_1} имеем

$$\chi_{E+E_1}(A) = \chi_E(A) + \chi_{E_1}(A).$$

Поэтому

$$\varphi(E+E_1) = \varphi(E) + \varphi(E_1).$$

Далее, рассмотрим n множеств E_i, $i = 1, 2, \ldots, n$, без общих точек. Обозначим через ε_i число, равное соответственно $+1$, -1, 0 в зависимости от того, будет ли $\varphi(E_i)$ больше нуля, меньше нуля или равно нулю, так что $\varepsilon_i \varphi(E_i) = |\varphi(E_i)|$.

По определению функции $\chi_E(A)$, функция $\Psi(A) = \sum_{i=1}^{n} \varepsilon_i \chi_{E_i}(A)$ принимает только значения $+1$, -1, 0 и $\|\Psi\| = 1$. Поэтому

$$|l(\Psi)| \leqslant \|l\| \, \|\Psi\| = \|l\|. \tag{8}$$

С другой стороны,

$$l(\Psi) = l\Big(\sum_{i=1}^{n} \varepsilon_i \chi_{E_i}\Big) = \sum_{i=1}^{n} \varepsilon_i l(\chi_{E_i}) = \sum_{i=1}^{n} \varepsilon_i \varphi(E_i) = \sum_{i=1}^{n} |\varphi(E_i)|. \tag{9}$$

Сравнивая (8) и (9), имеем

$$\sum_{i=1}^{n} |\varphi(E_i)| \leqslant \|l\|. \tag{10}$$

Таким образом $\varphi(E)$ есть абсолютно аддитивная функция множеств, причем

$$m(\varphi) \leqslant \|l\|. \tag{11}$$

Пусть тепер $F(A)$ есть произвольная непрерывная функция точки $A \subset M$ Построим последовательность разбиений S_k

$$M = E_1^{(k)} + E_2^{(k)} + \ldots + E_{n_k}^{(k)} \quad (k = 1, 2, \ldots),$$

где $E_i^{(k)}$ — измеримые множества попарно без общих точек; через $\delta(S_k)$ обозначим попрежнему наибольший из диаметров множеств $E_i^{(k)}$; будем предполагать, что $\delta(S_k) \to 0$.

Обозначим

$$F_k(A) = \sum_{i=1}^{n_k} F(A_i^{(k)}) \chi_{E_i^{(k)}}(A), \tag{12}$$

где $A_i^{(k)}$ — произвольная точка множества $E_i^{(k)}$. По определению функции $\chi_E(A)$,

$$F_k(A) = F(A_i^{(k)}), \quad \text{при } A \subset E_i^{(k)}.$$

Отсюда

$$|F(A)-F_k(A)|=|F(A)-F(A_i^{(k)})|\leqslant\eta^{(k)},$$

где через $\eta^{(k)}$ мы обозначили наибольшее из колебаний непрерывной функции $F(A)$ на множествах $E_i^{(k)}$, $i=1, 2, \ldots, n_k$.

Так как наибольший диаметр $\delta(S_k)$ множеств $E_i^{(k)}$ стремится к нулю, то и $\eta^{(k)}$ стремится к нулю. Но у нас [см. (1)]

$$\|F-F_k\|=\sup_{A\subset M}|F(A)-F_k(A)|\leqslant\eta^{(k)}.$$

Отсюда следует, что F_k равномерно стремится к F,

$$F_k\to F.$$

Поэтому

$$l(F_k)\to l(F).$$

Но

$$l(F_k)=l\left\{\sum_{i=1}^{n_k}F(A_i^{(k)})\chi_{E_i^{(k)}}(A)\right\}=\sum_{i=1}^{n_k}F(A_i^{(k)})\,l(\chi_{E_i^{(k)}})=\sum_{i=1}^{n_k}F(A_i^{(k)})\,\varphi(E_i^{(k)}).$$

Переходя к пределу по $k\to\infty$, получим

$$l(F)=\lim_{k\to\infty}l(F_k)=\int_M F\,d\varphi. \tag{13}$$

Таким образом мы доказали, что $l(F)$ выражается в форме (13). Остается вычислить $\|l\|$. По определению норма функционала $\|l\|$ есть наименьшая из постоянных K, для которых

$$|l(F)|\leqslant K\|F\|.$$

Но [см. неравенство (6)]

$$|l(F)|=\left|\int_M F\,d\varphi\right|\leqslant\|F\|\int_M|d\varphi|=\|F\|\,m(\varphi).$$

Отсюда

$$\|l\|\leqslant m(\varphi); \tag{14}$$

из (11) и (14) получаем

$$\|l\|=m(\varphi),$$

и теорема доказана полностью.

Положительно аддитивную (и аналогично, абсолютно аддитивную) функцию $\varphi(E)$ можно рассматривать как некоторую меру множества E, — естественное обобщение лебеговской меры (которая, очевидно, тоже положительно аддитивна). Целый ряд теорем, касающихся лебеговской меры, переносится на эту обобщенную меру. В качестве примера укажем на следующую теорему [аналогичную теореме Д. Ф. Егорова [1]].

Теорема 3. *Если последовательность непрерывных функций $f_n(A)$ сходится в каждой точке $A\subset M$ к некоторой функции $f(A)$, и $\varphi(E)$ — абсолютно аддитивная функция множеств, то для любого $\varepsilon>0$ существует множество $E_1\subset(E)$ такое, что $f_n(A)$ равномерно сходится к $f(A)$ на E_1 и $\varphi(M-E_1)\leqslant\varepsilon$.*

---

[1] См., например, Александров и Колмогоров, Введение в теорию функций действительного переменного, стр. 195.

Приведем следствие этой теоремы, которое нам понадобится.

Теорема 4. *Пусть $f_n(A)$ — последовательность непрерывных функций на M, ограниченная в совокупности:*

$$|f_n(A)| \leqslant K \text{ для всех } A \subset M \quad (n = 1, 2, 3, \ldots);$$

пусть $f_n(A)$ для каждого $A \subset M$ стремится к $f(A)$, где $f(A)$ — непрерывная функция на M. Тогда для любой абсолютно аддитивной функции множеств $\varphi(E)$ имеем

$$\lim_{n\to\infty} \int_M f_n(A)\, d\varphi = \int_M f(A)\, d\varphi.$$

Нам достаточно (в силу теоремы 1) ограничиться случаем положительной функции $\varphi(E)$. По теореме 3, для заданного $\varepsilon > 0$ можно выбрать такое E_1, что $\varphi(M - E_1) < \varepsilon$ и $f_n(A)$ равномерно сходится на E_1 к $f(A)$. При достаточно большом n будем иметь

$$|f_n(A) - f(A)| \leqslant \varepsilon \text{ на } E_1.$$

Далее, так как $f(A) = \lim\limits_{n\to\infty} f_n(A)$, то для всех $A \subset M$

$$|f(A)| \leqslant K;$$

поэтому

$$\left|\int_M [f_n(A) - f(A)]\, d\varphi\right| \leqslant$$
$$\leqslant \int_{E_1} |f_n(A) - f(A)|\, d\varphi \leqslant + \int_{M-E_1} |f_n(A) - f(A)|\, d\varphi \leqslant$$
$$\leqslant \varepsilon\varphi(M) + 2K\varphi(M - E_1) < \varepsilon[\varphi(M) + 2K].$$

В виду произвольности ε теорема доказана:

$$\int_M [f_n(A) - f(A)]\, d\varphi \to 0.$$

Заметим в заключение, что для абсолютно аддитивной функции множеств $\varphi(E)$ можно ввести не только понятие φ-измеримых множеств, но и понятие φ-измеримых функций $f(A)$ [$f(A)$ — функция точки A пространства M]. На эти φ-измеримые функции полностью переносится теория функций, измеримых по Лебегу, в частности, лебеговское интегрирование (так называемый интеграл Лебега-Стилтьеса-Радона). Заметим в частности, что (7) представляет собой также общий вид линейного функционала на L_M, но интеграл нужно понимать в смысле Лебега-Стилтьеса-Радона (см. Канторович и Фихтенгольц, Sur les opérations linéaires dans l'espace des fonctions bornées, Studia Mathematica, т. V. 1935).

§ 8. Слабая сходимость.

Во многих вопросах анализа встречается так называемая слабая сходимость элементов линейного пространства. Мы говорим, что последовательность элементов x_n пространства E *слабо сходится* к x и пишем

$$x_n \xrightarrow{\text{сл}} x,$$

если для любого линейного функционала $f \subset \overline{E}$ имеем

$$f(x_n) \to f(x).$$

Обычная сходимость

$$x_n \to x,$$

означающая стремление к нулю нормы $\| x_n - x \|$, называется иногда *сильной сходимостью*. В n-мерных пространствах сильная и слабая сходимости эквивалентны. В бесконечномерных сильная сходимость означает и слабую; обратное, вообще говоря, не верно.

Например, последовательность $y_n = \sin n\pi t$ в L_2, стремится слабо к $y = 0$. В самом деле, как мы уже указывали (см. § 6, стр. 111), всякий линейный функционал $f(y)$ в L_2 имеет вид

$$f(y) = \int_0^1 y(t)\varphi(t)\,dt,$$

где $\varphi(t)$ тоже входит в L_2. Имеем поэтому

$$f(y_n) = \int_0^1 \sin n\pi t\,\varphi(t)\,dt,$$

так что $f(y_n)$ совпадает с n-м коэфициентом Фурье для функции $\varphi(t)$ с интегрируемым квадратом. Следовательно, $\lim\limits_{n\to\infty} f(y_n) = 0$. Таким образом последовательность $y_n \equiv \sin n\pi t$ слабо сходится в L_2 к функции $y \equiv 0$. В то же время сильной сходимости y_n к $y \equiv 0$ нет, поскольку

$$\| y_n \| = \int_0^1 \sin^2 n\pi t\,dt = \frac{1}{2}.$$

Теорема 1 (Лебега). *Если последовательность $x_n \subset E$ слабо сходится к x то нормы $\| x_n \|$ ограниченны.*

Если $x_n \xrightarrow{\text{сл}} x$, то $(x_n - x) \xrightarrow{\text{сл}} \theta$. Мы можем поэтому ограничиться случаем $x_n \xrightarrow{\text{сл}} \theta$.

Допустим, что $x_n \xrightarrow{\text{сл}} \theta$ и нормы $\| x_n \|$ неограниченны.

Построим теперь подпоследовательность x_{n_i} последовательности x_n и последовательность функционалов $f_i \subset \overline{E}$ следующим образом: x_{n_1} есть произвольный элемент x_n, у которого $\| x_n \| > 1$. Пусть мы выбрали элементы x_{n_i} $(i = 1, 2, \ldots, k-1)$ и функционалы f_i $(i = 1, 2, \ldots, k-1)$. Мы выбираем тогда x_{n_k} так, чтобы выполнялись следующие два условия:

$$|f_i(x_{n_k})| < \frac{1}{4k}, \quad i = 1, 2, \ldots, k-1 \tag{1}$$

(так как для всякого линейного функционала $f(x_n) \to f(\theta) = 0$, то неравенства будут удовлетворяться, если взять n_k достаточно большим)

$$\| x_{n_k} \| > 3 \| x_{n_{k-1}} \|, \tag{2}$$

откуда $\| x_{n_k} \| > 3^k$ (в силу неограниченности $\| x_n \|$ можно при достаточно большом n_k найти x_{n_k}, удовлетворяющее этим неравенствам).

Данному x_{n_k} отнесем линейный функционал $f_k \subset \overline{E}$ такой, что

$$\|f_k\| = 1 \quad \text{и} \quad f_k(x_{n_k}) = \|x_{n_k}\|.$$

В силу теоремы 2, § 6, построить такой функционал f_k возможно.

Рассмотрим теперь линейный функционал

$$f(x) = \sum_{k=1}^{\infty} \frac{f_k(x)}{\|x_{n_k}\|}. \tag{3}$$

Так как последовательность норм

$$\sum_{k=1}^{\infty} \frac{\|f_k\|}{\|x_{n_k}\|} = \sum_{k=1}^{\infty} \frac{1}{\|x_{n_k}\|}$$

сходится ($\|x_{n_k}\| > 3^k$), то в силу полноты $\overline{E}$ правая часть формулы (3) сходится и дает в сумме линейный функционал $f(x)$.

Мы должны иметь (так как $x_{n_k} \xrightarrow{\text{сл}} \theta$)

$$f(x_{n_k}) \to f(\theta) = 0. \tag{4}$$

С другой стороны,

$$f(x_{n_k}) = \sum_{i=1}^{\infty} \frac{f_i(x_{n_k})}{\|x_{n_i}\|} = \sum_{i=1}^{k-1} \frac{f_i(x_{n_k})}{\|x_{n_i}\|} + \frac{f_k(x_{n_k})}{\|x_{n_k}\|} + \sum_{i=k+1}^{\infty} \frac{f_i(x_{n_k})}{\|x_{n_i}\|}. \tag{5}$$

При $i < k$ в силу (1) и $\|x_{n_i}\| \geqslant 1$, имеем

$$\frac{|f_i(x_{n_k})|}{\|x_{n_i}\|} \leqslant |f_i(x_{n_k})| \leqslant \frac{1}{4k}. \tag{6}$$

При $i > k$

$$\frac{|f_i(x_{n_k})|}{\|x_{n_i}\|} \leqslant \|f_i\| \frac{\|x_{n_k}\|}{\|x_{n_i}\|} \leqslant \frac{1}{3^{i-k}}. \tag{7}$$

Из (5), (6) и (7) следует

$$|f(x_{n_k})| \geqslant \frac{|f_k(x_{n_k})|}{\|x_{n_k}\|} - \sum_{i=1}^{k-1} \frac{|f_i(x_{n_k})|}{\|x_{n_i}\|} - \sum_{i=k+1}^{\infty} \frac{|f_i(x_{n_k})|}{\|x_{n_i}\|} \geqslant$$

$$\geqslant 1 - \frac{k-1}{4k} - \sum_{i=1}^{\infty} \frac{1}{3^i} > 1 - \frac{1}{4} - \frac{1}{2} = \frac{1}{4}. \tag{8}$$

Противоречие между соотношением (4) и неравенством (8) доказывает теорему.

Пример. Слабая сходимость в C_M.

Теорема 2. *Для того чтобы последовательность $y_n(A) \subset C_M$ слабо сходилась к $y(A) \subset C_M$, необходимо и достаточно, чтобы $y_n(A)$ ограниченно сходились к $y(A)$, т. е. чтобы последовательность y_n была равномерно ограничена и чтобы для любого $A \subset M$ $\lim\limits_{n\to\infty} y_n(A) = y(A)$.*

1. Условие необходимо. В число линейных функционалов $f(y)$ в C_M входят значения $y(A)$ для любого фиксированного $A \subset M$:

$$f(y) = y(A).$$

Условие слабой сходимости

$$f(y_n) \to f(y)$$

дает нам

$$\lim_{n \to \infty} y_n(A) = y(A) \tag{9}$$

для любого $A \subset M$.

Теорема 1 об ограниченности норм вместе с соотношением (9) и доказывает необходимость условия.

2. Условие достаточно. Достаточность условия теоремы вытекает из теоремы Рисса-Радона (§ 7). Каждый линейный функционал $f(y)$ в C_M имеет вид

$$f(y) = \int_M y(A)\, d\varphi,$$

где φ — абсолютно аддитивная функция множеств.

Если $y_n(A)$ на M ограниченно сходится к y, то для любой абсолютно аддитивной функции множеств φ (см. § 7, теорема 4)

$$\int_M y_n\, d\varphi \to \int_M y\, d\varphi,$$

следовательно, для любого линейного функционала $f(y_n) \to f(y)$, и достаточность доказана.

Нетрудно доказать следующие теоремы.

Для слабой сходимости в l_p *последовательности* $x^{(n)}(x_1^{(n)}, x_2^{(n)}, \ldots, x_k^{(n)}, \ldots)$ *к* $x(x_1, x_2, \ldots, x_k, \ldots)$ *необходимыми и достаточными условиями являются:*

1) *ограниченность норм* $\| x^{(n)} \|$,

2) *сходимость координат* $x_k^{(n)} \to x_k$, $n \to \infty$ *при любом* k.

Для слабой сходимости в L_p *последовательности функций* $y_n(t)$ *к* $y(t)$ *необходимыми и достаточными условиями являются:*

1) *ограниченность норм* $\| y_n \|$,

2) *сходимость неопределенных интегралов:*

$$\int_0^u y_n(t)\, dt \to \int_0^u y(t)\, dt \quad \textit{для} \quad 0 < u \leqslant 1.$$

Замечание. При условии $x_n \xrightarrow{\text{сл}} x$ мы всегда имеем

$$\| x \| \leqslant \overline{\lim} \| x_n \|.$$

В самом деле, пусть $\| x \| > \overline{\lim} \| x_n \|$. Тогда существует число c, такое, что

$$\| x \| > c > \overline{\lim} \| x_n \|.$$

Для всех x_n, начиная с некоторого места,

$$\| x_n \| < c.$$

С другой стороны, можно определить функционал f такой, что

$$|f(x)| = \|f\| \cdot \|x\| > c\|f\|.$$

Тогда имеем

$$|f(x_n)| \leqslant \|f\| \cdot \|x_n\| < c\|f\|$$

и $f(x_n)$ не стремится к $f(x)$, вопреки определению слабой сходимости.

Геометрическое определение слабой сходимости. Можно определить слабую сходимость в E следующим образом.

Рассмотрим совокупность (U) всех замкнутых выпуклых множеств U, заключающих бесчисленное множество точек последовательности x_n. Если пересечение всех $U \subset (U)$ не пусто, то оно состоит из единственной точки x — слабого предела последовательности x_n.

Это определение эквивалентно вышеприведенному.

Приведем еще одну интересную теорему.

Теорема 3. *Если* $x_n \xrightarrow{\text{сл}} x$ *в* E, *то* x *принадлежит к замкнутой линейной оболочке* $\overline{L}$ *последовательности* x_n. (*Иначе говоря,* x *есть сильный предел некоторой последовательности* $\sum\limits_{i=1}^{k_n} \lambda_i^{(n)} x_i$ *линейных комбинаций из членов последовательности* x_n.)

Допустим, что x не входит в $\overline{L}$. Можно построить линейный функционал f, равный нулю на $\overline{L}$ и не равный нулю в x: $f(x) = c \neq 0$ [1].

Для x_i, поскольку $x_i \subset \overline{L}$, будем иметь

$$f(x_i) = 0$$

и

$$\lim_{i \to \infty} f(x_i) = 0 \neq f(x),$$

что противоречит определению слабой сходимости.

Из теоремы 2 и теоремы 3 в применении к C (или C_M) следует:

Если последовательность $y_n(t)$ *непрерывных на* $[0, 1]$ (*на* M) *функций ограниченно сходится к непрерывной функции* $y(t)$, *то существует последовательность линейных комбинаций из* $y_n(t)$, *равномерно сходящаяся к* $y(t)$.

Слабая сходимость функционалов. Аналогично слабой сходимости элементов пространства можно определить и слабую сходимость функционалов. Пусть задана в $\overline{E}$ последовательность функционалов $f_n(x)$. Мы будем говорить, что f_n *слабо сходится*, если $f_n(x)$ сходится для любого $x \subset E$.

Обозначим

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

---

[1] Для этого достаточно функционал $f(x)$, тождественно равный нулю на $\overline{L}$, распространить на пространство $(\overline{L} + x)$ (см. § 6, доказательство теоремы Гана), положив

$$F(x') = f(x_1) + ct, \quad \text{где } x' = x_1 + tx, \ x_1 \subset \overline{L}, \ c \neq 0.$$

Тогда при $x_1 = \theta_E$ (всякая линейная оболочка, очевидно, содержит элемент θ_E) и $t = 1$, получим $F(x) = c$. Затем распространяем $F(x)$ произвольным образом на остальную часть пространства E.

Как уже выше было доказано (теорема 3, § 6), $f(x)$ есть также линейный функционал. $f(x)$ называется *слабым пределом* последовательности $f_n(x)$. Мы будем писать

$$f_n \overset{\text{сл}}{\to} f.$$

Сходимость

$$f_n \to f,$$

эквивалентную условию

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

будем называть *сильной сходимостью* функционалов.

Для сходимости функционалов f_n, определенных в полном пространстве E, имеет место теорема, аналогичная теореме 1:

Теорема 4. *Если $f_n \overset{\text{сл}}{\to} f$, то нормы $\| f_n \|$ ограничены.*

Доказательство ее совпадает в основном с доказательством теоремы 1.

Укажем еще следующую теорему:

Теорема 5. *Если пространство E сепарабельно, то единичная сфера $\| f \| \leqslant 1$ в $\overline{E}$ слабо компактна. Другими словами, из всякой последовательности $f_n \subset \overline{E}$, где $\| f_n \| \leqslant 1$, можно выбрать слабо сходящуюся подпоследовательность.*

В самом деле, пусть $x_1, x_2, \ldots, x_n, \ldots$ есть счетная всюду плотная сеть на E. Так как числа $|f_n(x_1)| \leqslant \| f_n \| \, \| x_1 \| \leqslant \| x_1 \|$ ограничены, то из последовательности f_n можно выбрать подпоследовательность T_1: $f_1', f_2', \ldots,$ для которой $f_n'(x_1)$ сходится. Из этой последовательности можно выбрать подпоследовательность T_2: $f_1'', f_2'', \ldots,$ для которой $f_n''(x_2)$ сходится, и т. д.

Получаем счетную систему вложенных друг в друга последовательностей $T_1 \supset T_2 \supset T_3 \supset \ldots \supset T_n \supset \ldots$, где $T_n = \{f_1^{(n)}, f_2^{(n)}, \ldots, f_k^{(n)}, \ldots\}$ и $f_k^{(n)}(x_n)$ при $k \to \infty$ сходится.

Построим диагональным процессом подпоследовательность

$$f_1^{(1)}, f_2^{(2)}, f_3^{(3)}, \ldots, f_n^{(n)}, \ldots$$

Она сходится для всех x_n нашей всюду плотной сети.

Возьмем произвольный элемент $x \subset E$. При любом $\varepsilon > 0$ можно найти для x элемент x_n нашей сети, для которого $\| x_n - x \| < \frac{\varepsilon}{3}$. Для любого $f \subset \overline{E}$

$$|f(x_n) - f(x)| = |f(x_n - x)| \leqslant \| f \| \, \| x_n - x \| \leqslant \| f \| \cdot \frac{\varepsilon}{3}.$$

В частности, так как $\| f_k^{(k)} \| \leqslant 1$, то

$$|f_k^{(k)}(x_n) - f_k^{(k)}(x)| \leqslant \frac{\varepsilon}{3}. \tag{10}$$

Вследствие сходимости последовательности $f_k^{(k)}(x_n)$ $(k \to \infty)$, при достаточно больших m и p

$$|f_m^{(m)}(x_n) - f_p^{(p)}(x_n)| < \frac{\varepsilon}{3}. \tag{11}$$

Пусть элементам x и x' из M отвечают в нашем отображении элементы $y\,(y_i)$ и $y'\,(y'_i)$ из m:

$$y_i = \rho(x, x_i) - \rho(x_0, x_i); \quad y'_i = \rho(x', x_i) - \rho(x_0, x_i).$$

Имеем

$$\| y - y' \|_m = \sup_i | y_i - y'_i | = \sup_i | \rho(x, x_i) - \rho(x', x_i) | \leqslant \rho(x, x'). \tag{3}$$

Пусть теперь ε — произвольное положительное число, меньшее $\rho(x, x')$. Существует точка x_n нашей всюду плотной сети на E, для которой

$$\rho(x, x_n) \leqslant \frac{\varepsilon}{2}.$$

Следовательно,

$$\rho(x', x_n) \geqslant \rho(x', x) - \rho(x, x_n) \geqslant \rho(x', x) - \frac{\varepsilon}{2}$$

и

$$| y_n - y'_n | = | \rho(x, x_n) - \rho(x', x_n) | \geqslant \rho(x', x_n) - \frac{\varepsilon}{2} \geqslant$$

$$\geqslant \rho(x', x) - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = \rho(x, x') - \varepsilon.$$

Отсюда

$$\| y - y' \|_m \geqslant \rho(x, x') - \varepsilon. \tag{4}$$

Сравнивая (3) и (4), получаем

$$\| y - y' \|_m = \rho(x, x').$$

Итак, расстояние между точками x, x' в M равно расстоянию между соответственными точками y, y' в m. Пространство M изометрично некоторой сепарабельной части M' пространства m. Линейная оболочка L над M' есть тоже сепарабельная часть m, которую можно рассматривать как сепарабельное пространство типа B.

Теорема 3. *Всякое сепарабельное пространство E типа B изометрично части C.*

Доказательство. Обозначим через U сферу $\| f \| \leqslant 1$ в $\overline{E}$, причем сходимость в U будем рассматривать как слабую сходимость функционалов f. U слабо компактно.

Можно рассматривать U как образ замкнутого множества P из отрезка $[0, 1]$; каждому $t \subset P$ отвечает функционал $f_t \subset U$, причем совокупность f_t совпадает с U и при $t_n \to t$, $f_{t_n} \overset{\text{сл}}{\to} f_t$. Выберем произвольный элемент $x \subset E$. По определению слабой сходимости функционалов

$$f_{t_n}(x) \underset{t_n \to t}{\overset{\text{сл}}{\to}} f_t(x)$$

При фиксированном x, $f_t(x)$ есть, следовательно, непрерывная функция от $t \subset P$, которую мы обозначим

$$f_t(x) = \varphi_x(t). \tag{5}$$

Функцию $\varphi_x(t)$, определенную на $P \subset [0, 1]$, можно определить на всем отрезке $[0, 1]$.

В самом деле, отрезок $[0, 1]$ состоит из P и счетного (или конечного) множества смежных интервалов (t_0, t_1), где $t_0; t_1 \subset P$. В концах t_0, t_1 интервала $[0, 1]$ функция $\varphi_x(t)$ определена; на самом интервале (t_0, t_1) будем считать $\varphi_x(t)$ линейной. Тем самым $\varphi_x(t)$ определена на всем $[0, 1]$.

При этом непрерывность $\varphi_x(t)$ сохранится, $\varphi_x(t) \subset C$.

По определению нормы в C

$$\|\varphi_x(t)\|_C = \max_{t \subset [0, 1]} |\varphi_x(t)|.$$

Но вследствие линейности $\varphi_x(t)$ на интервалах, смежных к P, максимум $\varphi_x(t)$ на $[0, 1]$ совпадает с максимумом $\varphi_x(t)$ на P:

$$\|\varphi_x(t)\|_C = \max_{t \subset P} |\varphi_x(t)|. \tag{6}$$

С другой стороны, в силу (5) для $t \subset P$

$$|\varphi_x(t)| = |f_t(x)| \leqslant \|f_t\| \cdot \|x\| \leqslant \|x\|.$$

Далее, для данного x можно построить функционал F нормы 1, для которого

$$\|F(x)\| = \|x\|_E.$$

Так как $F \subset U$, то существует $t_0 \subset P \subset [0, 1]$, для которого

$$f_{t_0} = F.$$

Следовательно,

$$|f_{t_0}(x)| = |\varphi_x(t_0)| = \|x\|_E. \tag{7}$$

Из (5) и (7) следует

$$\max_{t \subset [0, 1]} |\varphi_x(t)| = \|x\|_E. \tag{8}$$

Сравнивая (6) и (8), получаем

$$\|\varphi_x\|_C = \|x\|_E. \tag{9}$$

Каждому $x \subset E$ отвечает $\varphi_x \subset C$ той же нормы.

Из построения φ_x следует, что если $x \subset E$ отвечает $\varphi_x \subset C$ и $y \subset E$ отвечает $\varphi_y \subset C$, то $x - y$ отвечает

$$\varphi_{x-y} = \varphi_x - \varphi_y.$$

В силу формулы (9)

$$\|x - y\|_E = \|\varphi_x - \varphi_y\|_C,$$

что показывает, что мы имеем изометричное отображение E на часть C.

Из теорем 2 и 3 следует:

Теорема 4 (Банаха Мазура). *Всякое метрическое сепарабельное пространство изометрично некоторой части пространства C.*

§ 10. Приложения к вариационному исчислению.

Прежде чем приступить к изложению применений к вариационному исчислению, остановимся на некоторых теоремах относительно линейных операторов, на которых эти применения основаны.

Обратный оператор. Пусть оператор $y = s(x)$ $(y \subset E_1,\ x \subset E)$ отображает E на все пространство E_1. Будем при этом предполагать, что при $x \neq x_1$ также и $s(x) \neq s(x_1)$ (т. е. каждое y отвечает только *одному* x). При этом условии, обратно, каждому y отвечает единственное x, для которого $s(x) = y$.

Будем писать

$$x = s^{-1}(y). \tag{1}$$

Оператор s^{-1} будем называть *обратным* к s.

Теорема 1 (Банаха). *Пусть линейный оператор s, отображающий полное пространство E на полное же пространство E_1, имеет обратный оператор s^{-1}. Тогда s^{-1} есть тоже линейный оператор.*

Оператор

$$s(x) = y, \quad x \subset E, \quad y \subset E_1.$$

относит сумме элементов x сумму элементов y. Поэтому обратный оператор

$$x = s^{-1}(y)$$

относит сумме элементов y сумму элементов x:

$$s^{-1}(y_1 + y_2) = s^{-1}(y_1) + s^{-1}(y_2). \tag{2}$$

Аналогично показываем, что оператор s^{-1} обладает свойством однородности:

$$s^{-1}(\lambda y) = \lambda s^{-1}(y). \tag{3}$$

Нам остается доказать непрерывность оператора s^{-1}.

Обозначим через Y_n совокупность тех $y \subset E_1$, для которых

$$\| s^{-1}(y) \| \leqslant n \| y \|, \tag{4}$$

и через $X_n \subset E$ — совокупность соответственных x:

$$x = s^{-1}(y), \quad \text{где } y \subset Y_n.$$

Очевидно, для $x \subset X_n$

$$\| s(x) \| \geqslant \frac{1}{n} \| x \|. \tag{5}$$

Так как $E_1 = \sum_{n=1}^{\infty} Y_n$ и E_1 полно, то тем самым $E_1 = \sum_{n=1}^{\infty} \overline{Y}_n$.

На основании теоремы 3, § 2, все $\overline{Y}_n$ не могут быть нигде неплотными на E_1; одно из $\overline{Y}_n$ должно содержать сферу $S(y_0, r)$, на которой, таким образом, Y_n плотно. Очевидно, в этой сфере заключена некоторая меньшая сфера $\overline{S}(y_1, r_1)$, центр которой принадлежит уже Y_n.

Пусть y есть произвольный элемент с нормой r_1. Так как

$$y_1 + y \subset \overline{S}(y_1, r_1)$$

и на последней сфере Y_n плотно, то можно найти последовательность

$$y_1 + y^{(k)}$$

из Y_n, стремящуюся к $y_1 + y$, т. е.

$$y^{(k)} \to y;$$

при этом можно считать, что для всех k

$$r_1 \geqslant \| y^{(k)} \| \geqslant \frac{r_1}{2}. \tag{6}$$

Так как y_1 и $y_1 + y^{(k)}$ принадлежат Y_n, то по определению Y_n

$$\| s^{-1}(y_1) \| \leqslant n \| y_1 \|,$$

$$\| s^{-1}(y_1 + y^{(k)}) \| \leqslant n \| y_1 + y^{(k)} \| \leqslant n(\| y_1 \| + \| y^{(k)} \|) \leqslant n(\| y_1 \| + r_1)$$

[см. (6)]. Поэтому

$$\| s^{-1}(y^{(k)}) \| = \| s^{-1}(y_1 + y^{(k)}) - s^{-1}(y_1) \| \leqslant \| s^{-1}(y_1 + y^{(k)}) \| + \\ + \| s^{-1}(y_1) \| \leqslant n \| y_1 \| + n(\| y_1 \| + r_1) = n(2 \| y_1 \| + r_1). \tag{7}$$

Из (6) и (7) следует

$$\| s^{-1}(y^{(k)}) \| \leqslant N \frac{r_1}{2} \leqslant N \| y^{(k)} \|,$$

где

$$N = \frac{2n(2 \| y_1 \| + r_1)}{r_1}.$$

Итак, $y^{(k)}$ принадлежат Y_N.

Мы доказали, что всякий $y \subset E_1$ с нормой r_1 можно аппроксимировать элементами $y^{(k)}$ из Y_N. Для произвольного $y \subset E_1$ элемент $y' = \frac{r_1 y}{\| y \|}$ имеет норму r_1, поэтому $y' = \lim_{k \to \infty} y'_k$, $y'_k \subset Y_N$. Отсюда

$$y = \frac{\| y \|}{r_1} y' = \lim_{k \to \infty} \frac{\| y \|}{r_1} y'_k,$$

где $\frac{\| y \|}{r_1} y'_k$ в силу соотношений (3) и (4) также принадлежат Y_N. Итак, Y_N всюду плотно на E_1.

Рассмотрим теперь произвольный элемент y из E_1. Пусть $\| y \| = l$. Выделим последовательность элементов $y_1, y_2, \ldots, y_n, \ldots$ из Y_N, удовлетворяющую следующим условиям:

$$\left\| y - \sum_{i=1}^{n} y_i \right\| \leqslant \frac{l}{2^n}; \quad \| y_n \| \leqslant \frac{l}{2^{n-1}}. \tag{8}$$

Построить такую последовательность возможно. В самом деле, так как Y_N плотно на E_1, то существует элемент $y_1 \subset Y_N$ такой, что $\| y - y_1 \| \leqslant \frac{l}{2}$; $\| y_1 \| \leqslant l$. Далее, можно $y - y_1$ аппроксимировать таким y_2 из Y_N, что $\| (y - y_1) - y_2 \| = \| y - (y_1 + y_2) \| \leqslant \frac{l}{2^2}$, $\| y_2 \| \leqslant \| y - y_1 \| \leqslant \frac{l}{2}$, и т. д.

Мы имеем, очевидно,

$$y = \lim_{n \to \infty} \sum_{i=1}^{n} y_i. \tag{9}$$

Обозначая

$$x_i = s^{-1}(y_i), \tag{10}$$

имеем в силу (4) и (8)

$$\| x_i \| \leqslant N \| y_i \| \leqslant \frac{Nl}{2^{i-1}}. \tag{11}$$

Следовательно,

$$\left\| \sum_{i=1}^{n+m} x_i - \sum_{i=1}^{n} x_i \right\| = \left\| \sum_{i=n+1}^{n+m} x_i \right\| \leqslant \sum_{i=n+1}^{n+m} \| x_i \| < \frac{Nl}{2^n},$$

откуда явствует, что последовательность

$$\sum_{i=1}^{n} x_i$$

является последовательностью фундаментальной в E. Вследствие полноты E эта последовательность сходится к некоторому элементу x:

$$x = \lim_{n\to\infty} \sum_{i=1}^{n} x_i.$$

Так как оператор s непрерывен, то [см. (10) и (9)]

$$s(x) = \lim_{n\to\infty} \sum_{i=1}^{n} s(x_i) = \lim_{n\to\infty} \sum_{=1}^{n} y_i = y.$$

Поэтому, обратно,

$$x = s^{-1}(y).$$

Из (11) имеем

$$\|x\| = \lim_{n\to\infty} \|\sum_{i=1}^{n} x_n\| \leqslant 2Nl = 2N\|y\|.$$

Итак,

$$\|s^{-1}(y)\| \leqslant 2N\|y\|. \tag{12}$$

Неравенство (12), удовлетворяющееся для произвольно выбранного нами y, доказывает непрерывность оператора s^{-1}, что и завершает доказательство теоремы.

Пространство G. Рассмотрим два полных пространства E и E_1 и пусть имеем линейный оператор $y = s(x)$, отображающий E на все пространство E_1.

Каждому элементу $y \subset E_1$ отвечает некоторое множество $G_y \subset E$, состоящее из тех элементов $x \subset E$, для которых

$$s(x) = y\ ^{1)}.$$

В частности, нулевому элементу θ_{E_1} пространства E_1 отвечает линейное многообразие G_θ.

Если x_1 есть произвольный элемент из G_y, то G_y получаем путем „параллельного" сдвига G_θ на x_1; иными словами, для любого $x_0 \subset G_\theta$ имеем $x_0 + x_1 \subset G_y$, и обратно, если $x \subset G_y$, то $x - x_1 \subset G_\theta$.

В самом деле, если $x_1 \subset G_y$, $x_0 \subset G_\theta$, то $s(x_1) = y$, $s(x_0) = \theta$. Поэтому $s(x_1 + x_0) = y + \theta = y$, $x_1 + x_0 \subset G_y$.

Обратно, если x, $x_1 \subset G_y$, то $s(x) = s(x_1) = y$; $s(x - x_1) = s(x) - s(x_1) = 0$, $x - x_1 \subset G_\theta$.

Пример. Пусть мы имеем линейное отображение трехмерного пространства E_3 на плоскость E_2. Каждой точке плоскости E_2 отвечает при этом отображении прямая в пространстве E_3, причем все эти прямые параллельны между собой.

Обозначим через G совокупность всех G_y. G можно рассматривать как линейное пространство, если считать

$$G_y + G_{y_1} = G_{y+y_1}, \quad \lambda G_y = G_{\lambda y}.$$

---

[1]) См. Hausdorff, Zur Theorie der linearen metrischen Räume, Journ. f. reine. u. ang. Math., т. 167, 1932, стр. 294—311. Перевод этой статьи появится в качестве приложения к печатающемуся переводу книги Гаусдорфа, Теория множеств.

Определим теперь метрику в G следующим образом:

$$\| G_y \| = \inf_{x \subset G_y} \| x \|. \tag{13}$$

Легко убедиться, что введенная норма удовлетворяет всем необходимым аксиомам.

Полнота пространства G.

Из полноты E и E_1 легко вывести полноту G.

Теорема 2. *Если пространства E и E_1 полны, то и G есть полное пространство.*

Пусть задана фундаментальная последовательность G_{y_n} в G. Можно выделить из нее подпоследовательность G_{z_k}, где $z_k = y_{n_k}$, такую, что

$$\| G_{z_k} - G_{z_{k+1}} \| \leqslant \frac{1}{2^{k+1}}.$$

Определим теперь последовательность $x_1, x_2, \ldots, x_k, \ldots$ точек E, удовлетворяющую условиям $x_k \subset G_{z_k}$ и $\| x_k - x_{k+1} \| < \frac{1}{2^k}$.

Выбираем x_1 произвольно на G_{z_1}. Пусть $x_1, x_2, \ldots, x_k$ уже выбраны. Тогда в качестве x_{k+1} выбираем одну из точек $G_{z_{k+1}}$, для которой $\| x_{k+1} - x_k \| \leqslant \leqslant \frac{1}{2^k}$. Такая точка всегда существует. В самом деле, по определению метрики в G,

$$\inf_{\substack{x' \subset G_{z_k} \\ x'' \subset G_{z_{k+1}}}} \| x'' - x' \| = \| G_{z_{k+1}} - G_{z_k} \| \leqslant \frac{1}{2^{k+1}}.$$

Мы можем здесь $x' \subset G_{z_k}$ фиксировать, приняв его равным x_k [1]). Тогда

$$\inf_{x'' \subset G_{z_{k+1}}} \| x'' - x_k \| \leqslant \frac{1}{2^{k+1}}.$$

Поэтому существуют такие точки $x_{k+1} \subset G_{z_{k+1}}$, для которых

$$\| x_k - x_{k+1} \| \leqslant \frac{1}{2^k}.$$

Для определенной таким образом последовательности имеем

$$\| x_{k+m} - x_k \| = \Big\| \sum_{i=0}^{m-1} (x_{k+i+1} - x_{k+i}) \Big\| \leqslant$$
$$\leqslant \sum_{i=0}^{m-1} \| x_{k+i+1} - x_{k+i} \| \leqslant \sum_{i=0}^{m-1} \frac{1}{2^{k+i}} < \frac{1}{2^{k-1}}.$$

---

[1]) В самом деле, если, $x' \subset G_{z_k}$, то $x_k - x' \subset G_0$; поэтому, если $x'' \subset G_{z_{k+1}}$, то $x''' = x'' + (x_k - x') \subset G_{z_{k+1}}$, причем $\| x'' - x' \| = \| x''' - x_k \|$. Таким образом

$$\inf_{\substack{x' \subset G_{z_k} \\ x'' \subset G_{z_{k+1}}}} \| x'' - x' \| = \inf_{x''' \subset G_{z_{k+1}}} \| x''' - x_k \|.$$

Таким образом последовательность x_k есть фундаментальная последовательность в E, и так как E полное пространство, то x_k сходится. Обозначим

$$x = \lim_{n \to \infty} x_n. \tag{14}$$

Так как $x_n \subset G_{z_n}$, то $s(x_n) = z_n$ и в силу непрерывности s

$$z = s(x) = \lim_{n \to \infty} s(x_n) = \lim_{n \to \infty} z_n.$$

Но тогда и

$$G_{z_n} \to G_z.$$

В самом деле,

$$\| G_z - G_{z_n} \| = \inf_{\substack{x_n' \subset G_{z_n} \\ x' \subset G_z}} \| x_n' - x' \| \leqslant \| x_n - x \|,$$

а в силу (14) $\| x_n - x \| \to 0$.

Так как при любых x из G_y

$$\| y \| = \| s(x) \| \leqslant \| s \| \| x \|,$$

то, переходя в правой части к нижней грани, для всех $x \subset G_y$, получим

$$\| y \| \leqslant \| s \| \| G_y \|.$$

Отображение, относящее каждому G_y соответствующее y, есть отображение линейное:

$$y = s_1(G_y).$$

Оно имеет обратное отображение

$$G_y = s_1^{-1}(y).$$

Отсюда (в силу теоремы 1) заключаем: оператор s_1^{-1} тоже является линейным.

Теорема 3. *Пусть дан линейный оператор $y = s(x)$, отображающий E на E_1, и линейный функционал $f(x)$ в E. Если для $x \subset G_0$ имеем*

$$f(x) = 0,$$

то тогда в E_1 существует линейный функционал $l(y)$ такой, что

$$f(x) = l[s(x)]. \tag{15}$$

Эта теорема является обобщением следующего известного факта из теории линейных уравнений.

Даны k линейных форм в n-мерном пространстве $(x_1, x_2, \ldots, x_n)$ $(k \leqslant n)$:

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \quad 2, \ldots, k, \tag{16}$$

и форма

$$y_0 = \sum_{j=1}^{n} a_{0j} x_j.$$

Если из условий

$$y_i = \sum_{j=1}^{n} a_{ij} x_j = 0, \quad i = 1, 2, \ldots, k \tag{17}$$

следует

$$y_0 = \sum_{j=1}^{n} a_{0j} x_j = 0, \tag{18}$$

то

$$a_{0j} = \sum_{i=1}^{k} \lambda_i a_{ij},$$

и, значит,

$$y_0(x) = \sum_{i=1}^{k} \lambda_i y_i(x) \quad (\lambda \text{ — постоянные}). \tag{19}$$

Если рассматривать $y_1, y_2, \ldots, y_k$ как компоненты k-мерного вектора y, то формулы (16) дают нам оператор

$$y = y(x),$$

отображающий n-мерное пространство $(x_1, x_2, \ldots, x_n)$ на k-мерное пространство $(y_1, y_2, \ldots, y_k)$. Согласно нашей терминологии, совокупность векторов x, удовлетворяющих уравнению (17), образует многообразие $y(x) = \theta$. Обозначим $l(y) = \sum_{i=1}^{k} \lambda_i y_i$; очевидно $l(y)$ есть линейный функционал на $(y_1, y_2, \ldots, y_k)$. Тогда равенство (19) запишется в виде

$$y_0(x) = l[y(x)].$$

Приступим к доказательству нашей теоремы.

В условиях теоремы имеем

$$f(x') = f(x), \quad \text{если} \quad s(x') = s(x).$$

В самом деле, $s(x' - x) = s(x') - s(x) = \theta$, $x' - x \subset G_\theta$, и согласно условию

$$f(x' - x) = 0, \quad f(x') = f(x).$$

Обозначим через $\varphi(G_y)$ функционал, определяемый условием

$$\varphi(G_y) = f(x), \quad \text{если} \quad x \subset G_y. \tag{20}$$

Поскольку правая часть не зависит от выбора x на G_y, то этот функционал определен на G. Далее, из неравенства

$$|\varphi(G_y)| = |f(x)| \leqslant \|f\| \cdot \|x\|_{x \subset G_y},$$

переходя в правой части к нижней грани, получим

$$|\varphi(G_y)| \leqslant \|f\| \cdot \|G_y\|.$$

Таким образом $\varphi(G_y)$ есть линейный функционал на G_y.

С другой стороны, как мы видели,

$$G_y = s_1^{-1}(y).$$

Следовательно,

$$\varphi(G_y) = \varphi[s_1^{-1}(y)] = l(y), \tag{21}$$

где l — линейный функционал на E_1 (вследствие линейности функционала φ и оператора s_1^{-1}).

Так как для $x \subset G_y$, $s(x) = y$, то из (20) и (21) получаем

$$f(x) = l[s(x)],$$

и теорема доказана.

Перейдем теперь к применениям к вариационному исчислению.

Диференциал оператора. Пусть задан оператор s (вообще говоря нелинейный), отображающий E на E_1. Определим теперь диференциал $ds_a(h)$ оператора s в точке $a \subset E$ $(h \subset E)$. Существуют три определения диференциала:

1. *Диференциал Фреше:* $ds_a(h)$ есть главная линейная часть приращения $s(a+h) - s(a)$, именно

$$s(a+h) - s(a) = ds_a(h) + \varepsilon, \tag{22}$$

где $\|\varepsilon\| = \eta \|h\| \to 0$ при $\|h\| \to 0$, и $ds_a(h)$ — линейный оператор.

2. *Сильный диференциал:*

$$ds_a(h) = \lim_{t \to 0} \frac{s(a+th) - s(a)}{t} = \frac{d}{dt} s(a+th)\Big|_{t=0}, \tag{23}$$

причем предел понимается в смысле сильной сходимости.

2. *Слабый диференциал* определяется аналогично формулой (23), только предельный переход понимается в слабом смысле.

Для функционалов слабый и сильный диференциалы совпадают.

В случае существования диференциала Фреше, как можно непосредственно убедиться, существуют и (совпадающие с ним) слабый и сильный диференциалы. Из существования сильного диференциала следует и существование совпадающего с ним слабого, но вообще не следует существования диференциала Фреше.

Понятие вариации в классическом анализе совпадает у различных авторов с понятием диференциала в разных смыслах, чаще в смысле слабого или сильного диференциала.

В дальнейшем мы будем, не оговаривая это каждый раз, специально рассматривать операторы s, имеющие в каждой точке диференциал $ds_a(h)$ в смысле Фреше, причем будем считать, что $ds_a(h)$ непрерывен относительно обоих аргументов a и h, и притом в сильном смысле.

Уравнения в вариациях. Рассмотрим уравнение

$$s(x) = \theta' \tag{24}$$

(где θ' — нулевой элемент пространства E_1). Совокупность $x \subset E$, удовлетворяющих уравнению (24), будем называть *многообразием.* Если $s(a) = \theta'$, то совокупность $a + h \subset E$, где h удовлетворяет уравнению

$$ds_a(h) = \theta', \tag{25}$$

будем называть *линейным касательным многообразием,* а его уравнение (25) (по аналогии с теорией диференциальных уравнений) — *уравнением в вариациях* для многообразия (24).

Если линейный оператор $ds_a(h)$ отображает E на все пространство E_1, то точка a называется *обыкновенной.*

Имеет место теорема, являющаяся обобщением теоремы Пуанкаре для обычных уравнений в вариациях:

Теорема 4. *Если a — обыкновенная точка многообразия* (24), *то каждой точке $a + h$ линейного касательного многообразия* (25) *отвечает при* *достаточно малых* h *одна по крайней мере точка* $a+h'$ *многообразия* (24), *и обратно, каждой точке* $a+h'$ *многообразия* (24) *отвечает точка* $a+h$ *из* (25), *причем* $\|h'-h\|$ *есть величина высшего порядка малости сравнительно с* $\|h\|$ (*или с* $\|h'\|$).

Мы не будем приводить доказательства этой теоремы — его можно провести методом последовательных приближений аналогично доказательству соответственной теоремы для n-мерного случая.

Условие экстремума.

Теорема 5. *Если функционал* $f(x)$, *непрерывный и диференцируемый в* a, *достигает в точке* a *экстремума, то*

$$df_a(h)=0$$

(диференциал равен нулю для любых $h \subset E$).

В самом деле, точка a является, в частности, точкой экстремума f на любой прямой $a+th$ $(-\infty<t<\infty)$. Но на $a+th$, $f(a+th)$ есть функция параметра t. Так как при $t=0$ $f(a+th)$ достигает экстремума, то

$$\frac{d}{dt}f(a+th)\Big|_{t=0}=df_a(h)=0.$$

Рассмотрим теперь задачу *условного* экстремума. Пусть заданы на E оператор $s(x)$, отображающий E на E_1, и функционал $f(x)$.

Ищем экстремум $f(x)$ при условии

$$s(x)=\theta' \quad (\theta'=\theta_{E_1}).$$

Теорема 6. *Если экстремум функционала* $f(x)$ *на многообразии* (24) *достигается в обыкновенной точке* a *этого многообразия, то имеем*

$$df_a(h)=0$$

для всех $h \subset E$, *определяющих касательное линейное многообразие к* (24), *т. е. удовлетворяющих уравнению* (25):

$$ds_a(h)=\theta'.$$

В самом деле, пусть h принадлежит (25) и в то же время

$$df_a(h)=c\neq 0.$$

При любом t элементу $a+th$ отвечает в силу теоремы 4 элемент $a+th+h'$ многообразия (24), где $\|h'\|$ есть величина высшего порядка малости сравнительно с $\|th\|$, т. е. сравнительно с t. По определению диференциала

$$f(a+th+h')=f(a)+df_a(th+h')+\varepsilon=$$
$$=f(a)+t\,df_a(h)+df_a(h')+\varepsilon=f(a)+ct+df_a(h')+\varepsilon,$$

где ε есть величина высшего порядка малости сравнительно с t, и точно так же $df_a(h')$ вместе с $\|h'\|$ есть величина высшего порядка малости сравнительно с t.

При достаточно малом t знак разности $f(a+th+h')-f(a)$ совпадает со знаком ct, так что с переменой знака t и эта разность будет менять знак. Итак, в любой близости a найдутся элементы $a+th+h'$ многообразия (24), в которых

f принимает значения как бо́льшие, так и меньшие, чем в a; точка a не может быть точкой экстремума для $f(a)$ на (24). Теорема доказана.

Из теорем 3 и 6 следует:

Теорема 7. *Если в обыкновенной точке a многообразия* (24) *функционал $f(x)$ достигает экстремума, то существует линейный функционал l на E_1 такой, что*

$$df_a(h) = l[ds_a(h)]. \tag{26}$$

Обозначая

$$H(x) = f(x) - l[s(x)], \tag{27}$$

запишем условие (26) в виде

$$dH_a(h) \equiv 0. \tag{26'}$$

Задача нахождения условного экстремума для f при условии (26) решается как задача нахождения безусловного экстремума от $H(x)$ [1]).

Геометрически теорема 7 означает, что линейное касательное многообразие (25) к (24) заключено в линейном касательном многообразии к „поверхности уровня" функционала $f(x)$.

Можно сказать: в точке a условного экстремума многообразие (24), на котором ищется экстремум, касается „поверхности уровня" $f(x) = \text{const.}$ для функционала $f(x)$. Имеет место, следовательно, картина, совершенно аналогичная теории условного экстремума для функций конечного числа переменных.

Примеры. 1. *Изопериметрическая задача.* Ищется экстремум $f_0(x)$ при условии $f_i(x) = 0$, $i = 1, 2, \ldots, n$ (f_0, f_i — функционалы в E). Рассматривая f_i ($i = 1, 2, \ldots, n$) как компоненты n-мерного вектора f, обозначим через $\vec{f}(x)$ вектор с компонентами $f_i(x)$ ($i = 1, 2, \ldots, n$).

Пусть экстремум достигается в точке $x_0 \subset E$.

В силу теоремы 7 в n-мерном пространстве существует линейный функционал $l(\vec{f})$, такой, что для

$$H(x) = f_0(x) - l[\vec{f}(x)]$$

имеем

$$dH_{x_0} \equiv 0.$$

Но так как в n-мерном пространстве векторов $\vec{f}$

$$l(\vec{f}) = \sum_{i=1}^{n} \lambda_i f_i$$

(λ_i — постоянные), то

$$H(x) = f_0(x) - \sum_{i=1}^{n} \lambda_i f_i(x).$$

Получаем правило множителей Эйлера.

---

[1]) См. Л. А. Люстерник, Об условных экстремумах функционалов, Математический сборник, т. 41, вып. 3, 1934.

2. Если оператор s в условии (26) отображает E на C: $s(x) = y_x(t)$, где $y_x(t)$ — непрерывная функция аргумента $t \subset [0, 1]$, то по теореме Рисса выражение в формуле (27) будет иметь вид

$$H = f(x) - l[y_x(t)] = f(x) - \int_0^1 y_x(t)\, d\lambda(t),$$

где λ — функция с ограниченной вариацией.

Появляются функциональные „множители Лагранжа" $\lambda(t)$.

Пользуясь этим методом, можно было бы рассмотреть некоторые новые классы вариационных задач, например, изопериметрическую задачу со счетным числом переменных.