В самом деле, теорема 19 для числа m опирается на утверждение теоремы 18 только для $m-1$. С другой стороны, при доказательстве теоремы 18 для $m-1$ следует учесть, что функция $x \to B(u_1(x), u_2(x))$ является композицией функции $x \to (u_1(x), u_2(x))$ класса C^{m-1} и билинейной непрерывной, а, следовательно, бесконечно дифференцируемой функции $(u_1, u_2) \to B(u_1, u_2)$, и, значит, можно воспользоваться теоремой 19 для $m-1$.

Далее, теоремы 18 и 19 доказаны для $m=1$. Отсюда можно получить теорему 19 для $m=2$, а, следовательно, теорему 18 для $m=2$, затем теорему 19 для $m=3$, а, следовательно, теорему 18 для $m=3$, и т. д.

2°) Вычислять последовательные производные функции $h = g \circ f$ можно по явным формулам, но они довольно сложны. Так, например, если $\vec{X}$ и $\vec{Y}$ — векторы из $\vec{E}$, то

$$h''(x)(\vec{X}, \vec{Y}) = g''(f(x))(f'(x)\vec{X}, f'(x)\vec{Y}) + \\ + g'(f(x))(f''(x)(\vec{X}, \vec{Y})) \in \vec{G}. \qquad (III, 6; 38)$$

Если E, F, G — поля скаляров $\mathbb{K}$ и если выбрано $X = Y = 1$, то получается элементарная формула

$$h''(x) = g''(f(x)) f'^2(x) + g'(f(x)) f''(x). \qquad (III, 6; 39)$$

Иногда полезно иметь выражение производной порядка m для функции $h = g \circ f$ в частном случае $E = F = G = \mathbb{K}$. Мы его приведем в теореме 21_3.

§ 7. ФОРМУЛА ТЕЙЛОРА. МАКСИМУМ И МИНИМУМ

Так как производная $f^{(m)}(a)$ порядка m является m-линейным отображением, то можно вычислить ее значение на системе векторов $(\vec{X}, \vec{X}, \ldots, \vec{X}) \in \vec{E}^m$, определенной одним вектором $\vec{X}$ из $\vec{E}$. Выражение $f^{(m)}(a)(\vec{X}, \vec{X}, \ldots, \vec{X})$ удобно записывать кратко в виде

$$f^{(m)}(a)\vec{X}^m = f^{(m)}(a)(\vec{X}, \vec{X}, \ldots, \vec{X}). \qquad (III, 7; 1)$$

Имеется столько формул Тейлора, сколько существует формул конечных приращений.

Теореме 13A соответствует формула Тейлора, обобщающая теорему 6:

Теорема 20A. *Пусть f — вещественная m раз дифференцируемая функция, определенная на открытом множестве $\Omega \subset E$. Предположим, что замкнутый отрезок $[x, x+\vec{h}]$ пол-*

ностью лежит в Ω и что f имеет производную $(m+1)$-го порядка во всех точках открытого отрезка $]x, x+\vec{h}[$.

Тогда имеет место формула

$$f(x+\vec{h}) = f(x) + \frac{f'(x)}{1!}\vec{h} + \frac{f''(x)}{2!}\vec{h}^2 + \ldots$$
$$\ldots + \frac{f^{(m)}(x)}{m!}\vec{h}^m + \frac{f^{(m+1)}(x+\theta\vec{h})}{(m+1)!}\vec{h}^{m+1}, \quad \text{(III, 7; 1}_2\text{)}$$

где θ — вещественное число и $0 < \theta < 1$.

Доказательство. Обозначим через Φ вещественную функцию на $[0, 1]$, определенную формулой Φ: $t \to f(x+t\vec{h})$. По индукции проверяется, что на $[0, 1]$ существует производная порядка $k \leqslant m$ и на $]0, 1[$ — производная порядка $m+1$, причем

$$\Phi^{(k)}(t) = f^{(k)}(x+t\vec{h})\,\vec{h}^k. \quad \text{(III, 7; 1}_3\text{)}$$

Теперь достаточно к этой функции применить теорему 6 (формула Тейлора для вещественной функции от одной вещественной переменной) на отрезке $[0, 1]$, чтобы получить требуемый результат.

Далее, теореме 13 соответствует следующая формула Тейлора:

Теорема 20. *Пусть f есть m раз дифференцируемое отображение открытого множества $\Omega \subset E$ в F. Предположим, что замкнутый отрезок $[x, x+\vec{h}]$ полностью лежит в Ω и что f имеет производную порядка $m+1$ в каждой точке открытого отрезка $]x, x+\vec{h}[$, не превосходящую по норме числа M. Тогда имеет место формула*

$$\|\vec{A}\| = \left\| \overrightarrow{f(x+\vec{h}) - f(x)} - f'(x)\vec{h} - \frac{f''(x)}{2!}\vec{h}^2 - \ldots \right.$$
$$\left. \ldots - \frac{f^{(m)}(x)}{m!}\vec{h}^m \right\| \leqslant \frac{M\|\vec{h}\|^{m+1}}{(m+1)!}. \quad \text{(III, 7; 2)}$$

Доказательство. Это утверждение при $m = 0$ было уже доказано в теореме 13. Поэтому мы можем воспользоваться индукцией. Предположим, что это неравенство доказано для $m-1$, и докажем его справедливость для m, $m \geqslant 1$.

Рассмотрим функцию $\vec{g}$, определенную формулой

$$\overrightarrow{g(t)} = \overrightarrow{f(x+t\vec{h}) - f(x)} - f'(x)\,t\vec{h} - \ldots - \frac{f^{(m)}(x)}{m!}(t\vec{h})^m. \quad \text{(III, 7; 3)}$$

Это отображение отрезка $[0, 1]$ из $\mathbb{R}$ в $\vec{F}$, и $\vec{A} = \overrightarrow{g(1) - g(0)}$. Оно непрерывно на этом отрезке и, кроме того, имеет непрерывную производную на всем отрезке, определенную формулой

$$\overrightarrow{g'(t)} = f'(x + t\vec{h})\,\vec{h} - f'(x)\,\vec{h} - \ldots - \frac{t^{m-1}}{(m-1)!} f^{(m)}(x)\,\vec{h}^m. \quad \text{(III, 7; 4)}$$

Согласно 2-й формуле (III, 6; 19),

$$f^{(k)}(x)\,\vec{h}^k = \left((f')^{(k-1)}(x)\,\vec{h}^{k-1}\right) \cdot \vec{h}. \quad \text{(III, 7; 5)}$$

Поэтому (III, 7; 4) можно записать в виде

$$g'(t) = \left[f'(x + t\vec{h}) - f'(x) - \ldots - \frac{(f')^{(m-1)}}{(m-1)!}(x)\,(t\vec{h})^{m-1}\right]\vec{h} \in \vec{F}, \quad \text{(III, 7; 6)}$$

где выражение, стоящее в квадратных скобках, является элементом из $\mathscr{L}(\vec{E}; \vec{F})$. Будучи примененным к $\vec{h} \in \vec{E}$, оно дает элемент пространства $\vec{F}$.

Найдем теперь оценку нормы этой производной. Мы видим, что выражение в скобках является аналогом выражения $\vec{A}$, но относящимся к приращению $t\vec{h}$, порядку $m - 1$ и функции f', определенной на Ω, со значениями в $\mathscr{L}(\vec{E}; \vec{F})$. В силу индуктивного предположения, учитывая, что $(f')^{(m)} = f^{(m+1)}$, мы сможем написать следующую оценку, пригодную для $0 \leqslant t \leqslant 1$:

$$\|\overrightarrow{g'(t)}\| \leqslant M \frac{\|t\vec{h}\|^m}{m!} \|\vec{h}\| = M \|\vec{h}\|^{m+1} \frac{t^m}{m!}. \quad \text{(III, 7; 7)}$$

Пользуясь теперь леммой, приведенной в связи с теоремой 13 $\left(\text{заменяя там } f(x) \text{ на } \overrightarrow{g(x)} \text{ и } g(x) \text{ на } M\|\vec{h}\|^{m+1} \frac{x^{m+1}}{(m+1)!}\right)$, мы можем убедиться, что для функции $\vec{g}$ на отрезке $[0, 1]$ выполняется неравенство

$$\|\overrightarrow{g(1) - g(0)}\| \leqslant M \frac{\|\vec{h}\|^{m+1}}{(m+1)!}, \quad \text{(III, 7; 8)}$$

и тем самым справедлива формула (III, 7; 2).

Следствие. *Пусть L есть $(m+1)$-линейное непрерывное отображение $\vec{E}^{m+1}$ в $\vec{F}$. Тогда при условиях теоремы имеет ме-*

сто формула

$$\left\| \overrightarrow{f(x+\vec{h}) - f(x) - f'(x)\vec{h} - \frac{f''(x)}{2!}\vec{h}^2 - \ldots} \right.$$
$$\left. \ldots - \frac{f^{(m)}(x)}{m!}\vec{h}^m - \frac{L}{(m+1)!}\vec{h}^{m+1} \right\| \leqslant \omega \frac{\|\vec{h}\|^{m+1}}{(m+1)!}, \qquad \text{(III, 7; 9)}$$

где ω *является точной верхней гранью* $\|f^{(m+1)}(\xi) - L\|$, *когда* ξ *пробегает интервал* $]x, x+\vec{h}[$.

Доказательство. Утверждение легко сводится к случаю, когда L является $(m+1)$-линейной симметричной формой. Впрочем, это единственный случай, представляющий интерес. Затем теорему следует применить к функции $\xi \to f(\xi) - \frac{L}{(m+1)!}\overrightarrow{(\xi - x)}^{m+1}$, производные которой вычисляются по формуле (III, 6; 22_2).

В заключение заметим, что существует формула Тейлора, соответствующая определению производной (III, 3; 13), которую можно назвать формулой Тейлора для бесконечно малых приращений. Здесь только удобнее заменить $m+1$ на m.

Теорема 21. *Если f имеет в Ω производные порядка $\leqslant m-1$ и производную порядка m в точке a, то*

$$f(a+\vec{h}) = f(a) + f'(a)\vec{h} + \ldots + \frac{f^{(m)}(a)}{m!}\vec{h}^m + \vec{\alpha}\|\vec{h}\|^m, \qquad \text{(III, 7; 10)}$$

где $\vec{\alpha}$ *стремится к* $\vec{0}$ *вместе с* $\vec{h}$.

Будем доказывать эту формулу снова индукцией по m. Она верна для $m=1$. Предполагая, что она верна для целого $m-1$, докажем ее справедливость для $m \geqslant 2$. Функция

$$\overrightarrow{g(\vec{\xi})} = \overrightarrow{f(a+\vec{\xi}) - f(a) - f'(a)\vec{\xi} - \ldots - \frac{f^{(m)}(a)}{m!}\vec{\xi}^m}, \qquad \text{(III, 7; 11)}$$

определенная в некоторой окрестности точки $\vec{0}$ в $\vec{E}$, со значениями в $\vec{F}$ дифференцируема. Согласно (III, 6; 22_2), ее производная имеет вид

$$g'(\vec{\xi}) = f'(a+\vec{\xi}) - f'(a) - \ldots - \frac{(f')^{(m-1)}(a)}{(m-1)!}\vec{\xi}^{m-1} \in \mathscr{L}(\vec{E}; \vec{F}). \qquad \text{(III, 7; 12)}$$

Применим к g' формулу (III, 7; 10) для $m-1$:

$$g'(\xi) = \overrightarrow{\beta(\xi)} \|\vec{\xi}\|^{m-1}, \qquad \text{(III, 7; 12}_2\text{)}$$

где $\vec{\beta}(\vec{\xi})$ стремится к $\vec{0}$ вместе с $\vec{\xi}$.

Тогда формула конечных приращений (теорема 13) дает соотношения

$$\|\overrightarrow{g(\vec{h})}\| = \|\overrightarrow{g(\vec{h}) - g(\vec{0})}\| \leqslant \|\vec{h}\| \sup_{\vec{\xi} \in]\vec{0}, \vec{h}[} \|\overrightarrow{g'(\vec{\xi})}\| \leqslant$$

$$\leqslant \|\vec{h}\| \sup_{\vec{\xi} \in]\vec{0}, \vec{h}[} \|\overrightarrow{\beta(\xi)}\| \|\vec{\xi}\|^{m-1} \leqslant \varepsilon \|\vec{h}\|^m, \qquad \text{(III, 7; 12}_3\text{)}$$

где ε стремится к 0 вместе с $\|\vec{h}\|$, что и дает требуемый результат для m.

Применение формулы Тейлора для вычисления производных

Иногда легче получить разложение Тейлора, чем вычислить последовательные производные. Тогда эти производные можно получить из готового разложения. Другими словами, имеет место теорема, обратная к теореме 21:

Теорема 21₂. *Пусть f есть m раз дифференцируемое отображение открытого множества $\Omega \subset E$ в F. Если найдены k-линейные непрерывные симметричные отображения L_k пространств $\vec{E}^k$ в $\vec{F}$, $k = 1, 2, \ldots, m$, и некоторый элемент L_0 из F, такие, что*

$$f(a + \vec{h}) = L_0 + L_1\vec{h} + \frac{L_2}{2!}\vec{h}^2 + \ldots + \frac{L_m}{m!}\vec{h}^m + \vec{\alpha}\|\vec{h}\|^m, \qquad \text{(III, 7; 12}_4\text{)}$$

где $\vec{\alpha} = \vec{\alpha}(\vec{h})$ стремится к $\vec{0}$ вместе с $\vec{h}$, то всегда имеют место равенства

$$L_k = f^{(k)}(a), \quad k = 1, 2, \ldots, m, \quad \text{и} \quad L_0 = f(a).$$

Доказательство. Из теоремы 21 вытекает, что если мы положим

$$A_k = f^{(k)}(a) - L_k \in \mathscr{L}_k(\vec{E}^k; \vec{F}), \qquad \vec{A}_0 = \overrightarrow{f(a) - L_0} \in \vec{F}, \qquad \text{(III, 7; 12}_5\text{)}$$

то

$$\sum_{k=1}^{m} \frac{A_k}{k!}\vec{h}^k = \vec{\beta}\|\vec{h}\|^m, \qquad \text{(III, 7; 12}_6\text{)}$$

где $\vec{\beta}$ стремится к $\vec{0}$ вместе с $\vec{h}$.

Устремляя $\vec{h}$ к $\vec{0}$, мы увидим, что A_0 равно нулю. Предположим, что $A_0, A_1, A_2, \ldots, A_{k-1}$ — нули для $k \leqslant m$, и докажем, что тогда A_k также будет нулем. Отсюда будет вытекать, что все A_k, $k = 0, 1, 2, \ldots, m$, — нули, и теорема будет доказана.

Положим $\vec{h} = t\vec{X}$, где $\vec{X}$ фиксирован, и устремим скаляр t к 0. Тогда величина

$$t^k \left(\frac{A_k}{k!} \vec{X}^k + t \sum_{l=k+1}^{m} \frac{A_l}{l!} t^{l-k-1} \vec{X}^l \right) \qquad (\text{III}, 7; 12_7)$$

является бесконечно малой более высокого порядка, чем t^m, и тем более чем t^k, когда t стремится к 0. Это означает, что ее отношение к t^k стремится к 0, а, следовательно, $A_k\vec{X}^k = 0$ для любого $\vec{X} \in \vec{E}$. Доказательство по индукции будет закончено, если мы докажем такую лемму:

Лемма. *Пусть A есть k-линейное симметрическое отображение $\vec{E}^k$ в $\vec{F}$. Если для любого $\vec{X}$ из $\vec{E}$*

$$A\vec{X}^k = \vec{0} \in \vec{F}, \qquad (\text{III}, 7; 13)$$

то A равно нулю.

Эта лемма хорошо известна для $m = 2$, ибо тогда для любых $\vec{X} \in \vec{E}$ и $\vec{Y} \in \vec{E}$ в силу симметричности A имеем:

$$A(\vec{X} + \vec{Y}, \vec{X} + \vec{Y}) - A(\vec{X}, \vec{X}) - A(\vec{Y}, \vec{Y}) = 4A(\vec{X}, \vec{Y}),$$

и условие (III, 7; 13) влечет за собой обращение в нуль левой части равенства, а значит, и его правой части, откуда следует, что A равно нулю. Распространим это доказательство на общий случай.

Рассмотрим для фиксированных $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_k$ функцию h от $t = (t_1, t_2, \ldots, t_k) \in \mathbb{R}^k$:

$$h(t) = A(t_1\vec{X}_1 + t_2\vec{X}_2 + \ldots + t_k\vec{X}_k)^k. \qquad (\text{III}, 7; 13_2)$$

По условию она тождественно равна нулю. Эта функция является полиномом от $t_1, t_2, \ldots, t_k$. Следовательно, каждый коэффициент этого полинома равен нулю. В силу симметричности A коэффициент при $t_1, t_2, \ldots, t_k$ равен $k!A(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_k)$, откуда и следует утверждение леммы.

Замечание. Мы должны были предполагать функцию f дифференцируемой до порядка m. Это существенно. *Тот факт, что в окрестности a функция f имеет разложение вида*

(III, 7; 12_4), *вовсе не влечет за собой существование производных порядка* >1. Например, вещественная функция f вещественной переменной, определенная формулами $f(0)=0$, $f(x)=$ $=x^{m+1}\sin\frac{1}{x^{m+1}}$ для $x\neq 0$, имеет рассматриваемое разложение в начале координат с $L_k=0$. Она имеет производные всех порядков в дополнении к началу координат. Поскольку $|f(x)-f(0)|=|f(x)|\leqslant|x^{m+1}|$, в начале координат ее первая производная равна нулю. Но эта первая производная, равная $(m+1)x^m\sin\frac{1}{x^{m+1}}-\frac{m+1}{x}\cos\frac{1}{x^{m+1}}$, в точках $x\neq 0$ принимает сколь угодно большие значения в любой окрестности 0. Следовательно, она разрывна в начале координат и, значит, рассматриваемая функция не имеет там второй производной.

Применим теорему 21_2 к вычислению последовательных производных сложной функции.

Теорема 21_3. *Если g и f — дифференцируемые m раз скалярные функции скалярной переменной, то производная порядка m функции $h=g\circ f$ в точке a задается следующей формулой, в которой $f^{(p)}$ означает $f^{(p)}(a)$ и $g^{(q)}$ означает $g^{(q)}(f(a))$:*

$$h^{(m)}(a)=\sum_{k_1+2k_2+\ldots+mk_m=m}\frac{m!}{k_1!\,k_2!\,\ldots\,k_m!\,(1!)^{k_1}(2!)^{k_2}\ldots(m!)^{k_m}}\times$$
$$\times g^{(k_1+k_2+\ldots+k_m)}(f')^{k_1}(f'')^{k_2}\ldots(f^{(m)})^{k_m}. \qquad (\text{III}, 7; 13_3)$$

Доказательство. Воспользуемся разложением Тейлора

$$h(x)=g(f(x))=\sum_{0\leqslant l\leqslant m}\frac{(f(x)-f(a))^l}{l!}g^{(l)}+R_m. \qquad (\text{III}, 7; 13_4)$$

Так как

$$f(x)-f(a)=\sum_{1\leqslant k\leqslant m}\frac{(x-a)^k}{k!}f^{(k)}+\rho_m, \qquad (\text{III}, 7; 13_5)$$

то с точностью до членов высшего порядка малости по сравнению с $(x-a)^m$, когда x стремится к a, имеем:

$$(f(x)-f(a))^l\approx\left(\sum_{1\leqslant k\leqslant m}\frac{(x-a)^k}{k!}f^{(k)}\right)^l. \qquad (\text{III}, 7; 13_6)$$

Можно написать формулу возведения в степень суммы l слагаемых, обобщающую формулу бинома Ньютона (см. ниже формулу (III, 7; 26)). Согласно этой формуле,

$$(f(x)-f(a))^l=\sum_{k_1+k_2+\ldots+k_m=l}\frac{l!}{k_1!\,k_2!\ldots k_m!}\,\frac{(x-a)^{k_1+2k_2+\ldots+mk_m}}{(1!)^{k_1}(2!)^{k_2}\ldots(m!)^{k_m}}\times$$
$$\times(f')^{k_1}(f'')^{k_2}\ldots(f^{(m)})^{k_m}. \qquad (\text{III}, 7; 13_7)$$

По этой формуле вычисляется $\dfrac{(f(x)-f(a))^l}{l!}$, а затем, с точностью до бесконечно малых более высокого порядка, чем $(x-a)^m$, при x, стремящемся к a, находится выражение для $h(x)$:

$$h(x)\approx\sum_{\substack{0\leqslant l\leqslant m\\ k_1+k_2+\ldots+k_m=l}}\frac{(x-a)^{k_1+2k_2+\ldots+mk_m}}{k_1!\,k_2!\ldots k_m!\,(1!)^{k_1}(2!)^{k_2}\ldots(m!)^{k_m}}\times$$
$$\times(f')^{k_1}(f'')^{k_2}\ldots(f^{(m)})^{k_m}g^{(l)}. \qquad (\text{III}, 7; 13_8)$$

Поскольку известно, что h дифференцируема m раз (теорема 19), величина $h^{(m)}(a)$ должна равняться произведению $m!$ и коэффициента при $(x-a)^m$ в предыдущем разложении. Отсюда и вытекает высказанное утверждение.

На этом примере видно, какую большую роль может сыграть разложение Тейлора в вычислении производных порядка >1 *при условии, что существование искомых производных доказано другими путями.*

Формула Тейлора относительно некоторой системы координат

Теорема 21_4. *Если пространство E конечномерно, в нем выбрана система координат и вектор $\vec{h}$ имеет координаты h_1, h_2, ..., h_n, то, пользуясь обозначениями* (III, 6; 26—27), *выражение* $\sum\limits_{k=0}^{m}\dfrac{f^{(k)}(x)}{k!}\vec{h}^k$, *входящее в правую часть формулы Тейлора, можно записать в виде*

$$\sum_{k=0}^{m}\frac{f^{(k)}(x)}{k!}\vec{h}^k=\sum_{|\vec{p}|\leqslant m}\frac{\overrightarrow{D^{\vec{p}}f(x)}}{\vec{p}!}\vec{h}^{\vec{p}}\in F. \qquad (\text{III}, 7; 14)$$

Доказательство. Согласно формуле (III, 6; 25),

$$\frac{f^{(k)}(x)}{k!}\vec{h}^k=\frac{1}{k!}\sum_{j_1,j_2,\ldots,j_k}\frac{\overrightarrow{\partial^k f}}{\partial x_{j_1}\partial x_{j_2}\ldots\partial x_{j_k}}(x)\,h_{j_1}h_{j_2}\ldots h_{j_k}\in\vec{F}\,^{1)}$$
$$(\text{III}, 7; 14_2)$$

---

[1]) В формулу (III, 6; 25) входили векторы $\vec{X}_1$, $\vec{X}_2$, ..., которые здесь равны $\vec{h}$.

В этой формуле $j_1, j_2, \ldots, j_k$ — индексы, принимающие независимо друг от друга всевозможные значения из множества $\{1, 2, \ldots, n\}$. Объединим между собой все конечные последовательности $j_1, j_2, \ldots, j_k$, у которых p_1 индексов равны 1, p_2 индексов равны 2, ..., p_n индексов равны n. В сумме (III, 7; 14$_2$) они дадут элемент, который запишется в обозначениях (III, 6; 25 и 27) в виде

$$\overrightarrow{\frac{D^{\vec{p}} \vec{f}(x) \vec{h}^{\vec{p}}}{k!}} = \overrightarrow{\left(\frac{\partial}{\partial x_1}\right)^{p_1} \left(\frac{\partial}{\partial x_2}\right)^{p_2} \cdots \left(\frac{\partial}{\partial x_n}\right)^{p_n} f(x)} \frac{h_1^{p_1} h_2^{p_2} \ldots h_n^{p_n}}{k!} \in \vec{F}. \quad \text{(III, 7; 14}_3\text{)}$$

Членов этого вида имеется столько, сколько существует отображений j множества $\{1, 2, \ldots, k\}$ в множество $\{1, 2, \ldots, n\}$, где $\{1\}$ имеет прообраз, состоящий из p_1 элементов, $\{2\}$ — из p_2 элементов, ..., $\{n\}$ — из p_n элементов, $|\vec{p}| = p_1 + p_2 + \ldots \ldots + p_n = k$. Это число $\gamma_{p_1, p_2, \ldots, p_n} = \gamma_{\vec{p}}$ называется числом «перестановок с повторениями» n объектов $\{1, 2, \ldots, n\}$, где объект 1 берется p_1 раз, объект 2 берется p_2 раз, ..., объект n берется p_n раз. Существуют разные способы вычисления этого числа. Для того чтобы определить интересующее нас отображение j, нам надо сначала выбрать часть из $\{1, 2, \ldots, k\}$, состоящую из p_1 элементов, которая и будет составлять $j^{-1}\{1\}$. Число возможных выборов равно числу подмножеств, содержащих по p_1 элементов из $\{1, 2, \ldots, k\}$, т. е. $\frac{k!}{(p_1)!\,(k-p_1)!}$. Сделав этот выбор, надо в дополнении $\{1, 2, \ldots, k\} - j^{-1}\{1\}$, состоящем из $k - p_1$ элементов, выбрать произвольно p_2 элементов, составляющих $j^{-1}\{2\}$. При этом число возможных выборов равно $\frac{(k-p_1)!}{p_2!\,(k-p_1-p_2)!}$. Так можно продолжать далее. Число таких отображений j равно

$$\gamma_{\vec{p}} = \frac{k!}{p_1!\,(k-p_1)!} \cdot \frac{(k-p_1)!}{p_2!\,(k-p_1-p_2)!} \cdots \frac{(k-p_1-p_2-\ldots-p_{n-1})!}{p_n!\,(k-p_1-p_2-\ldots-p_n)!} =$$
$$= \frac{k!}{p_1!\,p_2! \ldots p_n!} = \frac{k!}{\vec{p}!}, \quad \text{(III, 7; 15)}$$

а, следовательно,

$$\frac{\gamma_{\vec{p}}}{k!} = \frac{1}{\vec{p}!}, \quad \text{(III, 7; 16)}$$

и теорема доказана.

В *принятых обозначениях полученная формула совпадает с формулой, соответствующей функциям одной переменной.* В левой части формулы Тейлора $f(x + \vec{h})$ лежит в аффинном пространстве F, все члены правой ее части принадлежат векторному пространству $\vec{F}$, кроме первого $f(x)$, соответствующего $\vec{p}=\vec{0}$ и лежащего в аффинном пространстве F.

Следствие. *Если при выполнении условий теоремы найдены такие элементы* $\vec{c}_{\vec{p}}$ $\left(\vec{c}_{\vec{p}}\in\vec{F}\right.$ *для* $\vec{p}\neq\vec{0}$, $\left.c_{\vec{0}}\in F\right)$, $|\vec{p}|\leqslant m$, *что*

$$f(a+\vec{h})=\sum_{|\vec{p}|\leqslant m}\vec{c}_{\vec{p}}\frac{\vec{h}^{\vec{p}}}{\vec{p}!}+\vec{\alpha}\|\vec{h}\|^m, \tag{III,7;17}$$

где $\vec{\alpha}\in\vec{F}$ *стремится к* $\vec{0}$ *вместе с* $\vec{h}$, *и если известно, что* f *дифференцируема* m *раз в* a, *то имеют место равенства*

$$\vec{c}_{\vec{p}}=\overrightarrow{D^{\vec{p}}f}(a). \tag{III,7;18}$$

Доказательство. Положим $\vec{a}_{\vec{p}}=\overrightarrow{D^{\vec{p}}f}(a)-\vec{c}_{\vec{p}}$ (здесь также $\vec{a}_{\vec{0}}\in\vec{F}$). Тогда, составляя разность, можно убедиться, что норма «полинома»[1]) (с векторными коэффициентами) относительно $h_1, h_2, \ldots, h_n$:

$$\sum_{|\vec{p}|\leqslant m}\vec{a}_{\vec{p}}\vec{h}^{\vec{p}}=\sum\vec{a}_{p_1,p_2,\ldots,p_n}h_1^{p_1}h_2^{p_2}\ldots h_n^{p_n} \tag{III,7;19}$$

является бесконечно малой более высокого порядка, чем $\|\vec{h}\|^m$, при $\vec{h}$, стремящемся к $\vec{0}$. Полагая $\vec{h}=t\vec{X}$ и устремляя t к 0, с помощью индукции, аналогичной той, которая применялась при доказательстве теоремы 21_2 (но без использования леммы), можно убедиться в том, что однородные части полинома $\sum\limits_{|\vec{p}|\leqslant m}\vec{a}_{\vec{p}}\vec{X}^{\vec{p}}$ степеней 0, 1, 2, ..., m тождественно равны нулю. Но у тождественно равного нулю полинома все коэффициенты равны нулю[2]), чем и заканчивается доказательство следствия.

Упражнение. Применяя следствие, получите новое доказательство формулы Лейбница (III,6;36) в духе теоремы 21_3.

Теорема 22. *Для того чтобы отображение* f *открытого связного множества* $\Omega\subset E$ *в пространство* F *имело производ-*

[1]) См. далее замечание о полиномах к теореме 22.

[2]) Это обычно доказывается при помощи индукции по числу n переменных $X_1, X_2, \ldots, X_n$ только для полиномов со скалярными коэффициентами. Но этот результат остается верным для векторных коэффициентов, причем сохраняется то же самое доказательство.

ную порядка $m+1$, равную нулю в Ω, необходимо и достаточно, чтобы отображение f было полиномом степени $\leqslant m$.

Доказательство. Пусть сначала $m=0$. Мы уже видели, что производная постоянной равна нулю, даже если Ω не связно. Нам надо доказать, что в случае связного множества Ω отображение f множества Ω в F, имеющее всюду равную нулю производную, постоянно.

Пусть x — произвольная точка Ω. Существует такое число $\rho_x>0$, что открытый шар B_x с центром в x радиуса ρ_x целиком лежит в Ω. При $\|\vec{h}\|<\rho_x$ можно на отрезке $[x; x+\vec{h}]$, лежащем в Ω, применить теорему о конечных приращениях и получить неравенство

$$\|\overrightarrow{f(x+\vec{h})-f(x)}\|\leqslant \sup_{\vec{\xi}\in\,]x,\,x+\vec{h}[}\|f'(\xi)\|=0, \tag{III,7;20}$$

из которого следует, что $f(x)$ постоянна в B_x.

Пусть теперь a — произвольная фиксированная точка множества Ω. Обозначим через A множество точек ξ из Ω, таких, что $f(\xi)=f(a)$. Так как A является прообразом множества $\{f(a)\}\in F$ при непрерывном отображении f, то оно замкнуто. Это множество также и открыто, ибо для $x\in A$ любой указанный выше шар принадлежит A. Поскольку Ω связно и A непусто, A совпадает со всем Ω, что и доказывает теорему для $m=0$.

Предположим теперь, что $m\geqslant 1$. Прежде всего надо дать определение полинома на E со значениями в F. Мы ограничимся частным случаем, когда E n-мерно.

Если выбрана система координат $0, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, то полином относительно $x_1, x_2, \ldots, x_n$ степени $\leqslant n$ со значениями в F определяется как функция вида

$$f(x)=\sum_{|\vec{p}|\leqslant m}\vec{a}_{\vec{p}}x^{\vec{p}}=\sum_{p_1+p_2+\ldots+p_n\leqslant m}a_{p_1,p_2,\ldots,p_n}x_1^{p_1}x_2^{p_2}\ldots x_n^{p_n}, \tag{III,7;21}$$

где $a_{\vec{p}}$ принадлежат $\vec{F}$, кроме $a_{\vec{0}}$, лежащего в F.

Эта функция имеет тот же вид в любой другой системе координат и называется полиномом на E степени $\leqslant m$ со значениями в F. Ее производная порядка $m+1$ равна нулю независимо от того, связно или нет множество Ω, поскольку, согласно (III, 7; 21), ее частные производные порядка $m+1$ равны нулю.

Обратно, пусть Ω связно, и пусть f — отображение множества Ω в пространство F, $(m+1)$-я производная которого равна нулю. Пусть $a \in \Omega$. Формула Тейлора порядка m относительно любой системы координат имеет нулевой остаточный член $\vec{R}_m$ и, значит, отображение f в указанном шаре B_a является полиномом P степени $\leqslant m$. Обозначим теперь через k такое целое число, что $(f-P)^{(k)} \equiv 0$ в Ω. Такое целое k существует, например $k=m+1$. Тогда, если $k \geqslant 1$, то $(f-P)^{(k-1)}$ является функцией, первая производная которой равна нулю в связном множестве Ω. Значит, она постоянна в Ω. Но в B_a она равна нулю и, следовательно, она равна нулю в Ω. Таким образом, $k-1$ обладает тем же свойством, что и k. Шаг за шагом спускаемся до $k=0$ и получаем $f \equiv P$ в Ω, что и завершает доказательство теоремы.

Замечания. 1°) Конечно, это заключение не верно, если множество Ω не связно. Достаточно взять в качестве Ω дополнение к началу координат на вещественной прямой $\mathbb{R}$ и положить $m=0$. Вещественнозначная функция, определенная на Ω, имеющая всюду равную нулю производную, не обязательно постоянна в Ω. Она может быть равной некоторой постоянной на полупрямой $]-\infty, 0[$ и другой постоянной на полупрямой $]0, +\infty[$.

2°) Предположим, что у нас нет формулы (III, 7; 15), определяющей $\gamma_{\vec{p}}$. Ее можно очень просто получить следующим образом.

Не считая $\gamma_{\vec{p}}$ известным, напишем формулу Маклорена для полинома степени m:

$$\overrightarrow{f(x)} = \sum_{|\vec{p}| \leqslant m} \gamma_{\vec{p}} \frac{\overrightarrow{D^{\vec{p}} f(0)}}{k!} \vec{x}^{\vec{p}} = \sum_{|\vec{p}| \leqslant m} c_{\vec{p}} \frac{\overrightarrow{D^{\vec{p}} f(0)}}{\vec{p}!} \vec{x}^{\vec{p}}, \qquad \text{(III, 7; 22)}$$

где $c_{\vec{p}} = \frac{\vec{p}!}{k!} \gamma_{\vec{p}}$, и покажем, что $c_{\vec{p}} = 1$.

Пусть $\vec{q} \in \mathbb{N}^n$, $\vec{q} \leqslant \vec{p}$. Тогда

$$D^{\vec{q}} \frac{\vec{x}^{\vec{p}}}{\vec{p}!} = \left(\left(\frac{\partial}{\partial x_1}\right)^{q_1} \frac{x_1^{p_1}}{p_1!}\right)\left(\left(\frac{\partial}{\partial x_2}\right)^{q_2} \frac{x_2^{p_2}}{p_2!}\right) \cdots \left(\left(\frac{\partial}{\partial x_n}\right)^{q_n} \frac{x_n^{p_n}}{p_n!}\right) =$$

$$= \begin{cases} \dfrac{\vec{x}^{\vec{p}-\vec{q}}}{(\vec{p}-\vec{q})!}, & \text{если } \vec{q} \leqslant \vec{p}\ ^{1)}, \\ 0, & \text{если } \vec{q} \nleqslant \vec{p}. \end{cases} \qquad \text{(III, 7; 23)}$$

[1]) Отношение порядка $\vec{q} \leqslant \vec{p}$ в $\mathbb{N}^n$ является частичным, а не полным: $\vec{q} \nleqslant \vec{p}$ вовсе не означает, что $\vec{q} > \vec{p}$. Так, например, при $n=2$ мы имеем $(1, 1) \nleqslant (0, 2)$, но при этом не выполняется и неравенство $(1, 1) > (0, 2)$.

Теперь дифференцирование $D^{\vec{q}}$ выражения (III, 7; 22) для $|\vec{q}|\leqslant m$ дает:

$$\overrightarrow{D^{\vec{q}}f(x)}=\sum_{\substack{|\vec{p}|\leqslant m\\ \vec{p}\geqslant\vec{q}}} c_{\vec{p}}\frac{\overrightarrow{D^{\vec{p}}f(0)}}{(\vec{p}-\vec{q})!}\vec{x}^{\vec{p}-\vec{q}}. \qquad \text{(III, 7; 24)}$$

Положим $\vec{x}=\vec{0}$. Тогда $\vec{x}^{\vec{p}-\vec{q}}=0$ для $\vec{p}\neq\vec{q}$ и $\vec{x}^{\vec{p}-\vec{q}}=1$ для $\vec{p}=\vec{q}$. Поэтому

$$\overrightarrow{D^{\vec{q}}f(0)}=c_{\vec{q}}\overrightarrow{D^{\vec{q}}f(0)}, \qquad \text{(III, 7; 25)}$$

откуда $c_{\vec{q}}=1$, что и требовалось доказать.

Положим, в частности, $f(x)=(x_1+x_2+\ldots+x_n)^k$. Теорема о сложной функции говорит о том, что любая частная производная порядка l этой функции равна $k(k-1)\ldots$ $\ldots(k-l+1)\,(x_1+x_2+\ldots+x_n)^{k-l}$. При $x_1=x_2=\ldots$ $\ldots=x_n=0$ эта производная равна нулю, кроме случая $l=k$, когда она равна $k!$. Поэтому формула Маклорена дает:

$$(x_1+x_2+\ldots+x_n)^k=\sum_{|\vec{p}|=k}\gamma_{\vec{p}}\vec{x}^{\vec{p}}=\sum_{|\vec{p}|=k}\frac{k!}{\vec{p}!}\vec{x}^{\vec{p}}. \qquad \text{(III, 7; 26)}$$

Итак, $\gamma_{p_1,p_2,\ldots,p_n}=\dfrac{(p_1+p_2+\ldots+p_n)!}{p_1!\,p_2!\ldots p_n!}$ *есть коэффициент при* $x_1^{p_1}x_2^{p_2}\ldots x_n^{p_n}$ *в разложении* $(x_1+x_2+\ldots+x_n)^{p_1+p_2+\ldots+p_n}$, *что является обобщением формулы бинома Ньютона.* Это, впрочем, очевидно, если воспользоваться комбинаторным определением $\gamma_{\vec{p}}$, данным после (III, 7; 14).

Коэффициенты бинома имеют очень простой вид, что позволяет сразу написать формулу Маклорена для функции двух скалярных переменных x и y. Полагая

$$\vec{p}=\frac{\partial\vec{f}}{\partial x}(0,0),\quad \vec{q}=\frac{\partial\vec{f}}{\partial y}(0,0),$$

$$\vec{r}=\frac{\overrightarrow{\partial^2 f}}{\partial x^2}(0,0),\quad \vec{s}=\frac{\overrightarrow{\partial^2 f}}{\partial x\,\partial y}(0,0),\quad \vec{t}=\frac{\overrightarrow{\partial^2 f}}{\partial y^2}(0,0), \qquad \text{(III, 7; 26}_2\text{)}$$

$$\vec{\alpha}=\frac{\overrightarrow{\partial^3 f}}{\partial x^3}(0,0),\quad \vec{\beta}=\frac{\overrightarrow{\partial^3 f}}{\partial x^2\,\partial y}(0,0),\quad \vec{\gamma}=\frac{\overrightarrow{\partial^3 f}}{\partial x\,\partial y^2}(0,0),\quad \vec{\delta}=\frac{\overrightarrow{\partial^3 f}}{\partial y^3}(0,0),$$

получим:

$$f(x,\ y)=f(0,\ 0)+\vec{p}x+\vec{q}y+\frac{1}{2}(\vec{r}x^2+2\vec{s}xy+\vec{t}y^2)+$$
$$+\frac{1}{6}(\vec{\alpha}x^3+3\vec{\beta}x^2y+3\vec{\gamma}xy^2+\vec{\delta}y^3)+\ \ldots \qquad (\mathrm{III}, 7; 26_3)$$

Применение к изучению максимумов и минимумов. Определения

Пусть f — *вещественнозначная* функция, определенная на топологическом пространстве E. Говорят, что f имеет в точке $a\in E$ *относительный максимум*[1]), если существует некоторая окрестность $\mathscr{V}$ точки $a\in E$, такая, что сужение f на эту окрестность имеет в точке a максимум, т. е. если для любой точки x из $\mathscr{V}$ имеет место неравенство $f(x)\leqslant f(a)$. Относительный максимум называют *строгим*, если можно выбрать $\mathscr{V}$ таким образом, чтобы для $x\neq a$ из $\mathscr{V}$ имело место строгое неравенство $f(x)<f(a)$.

Аналогичное определение дается для *относительного минимума и строгого относительного минимума.*

Конечно, максимум функции f, т. е. ее абсолютный максимум, является одновременно и относительным максимумом, в то время как противоположное утверждение, вообще говоря, не верно.

Говорят, что *функция f имеет в точке a экстремум*, если она имеет в этой точке относительный максимум или относительный минимум.

Необходимые условия экстремума

Теорема 23. *Пусть f — вещественнозначная функция, определенная на открытом множестве Ω аффинного нормированного пространства E и дифференцируемая на нем. Необходимым условием существования в точке $a\in\Omega$ относительного максимума или минимума является обращение в нуль производного отображения $f'(a)\in\mathscr{L}(\vec{E};\mathbb{R})=\overleftarrow{E}$.*

Доказательство. Заметим прежде всего, что если E конечномерно и в нем выбрана система координат 0, $(\vec{e}_i)_{i\in I}$, то это необходимое услов\[е выражается в общеизвестной форме: *для того чтобы f в точке a имела относительный максимум или минимум, необходимо, чтобы частные производные $\partial f/\partial x_i$*

[1]) Вместо того чтобы говорить, что функция f имеет в точке a максимум или относительный максимум и т. д., иногда говорят, что точка a является для f максимумом или относительным максимумом и т. д. Это не верно, так как максимум равен $f(a)$, а не a!

в точке a были равны нулю или чтобы дифференциал в точке a: $\sum\limits_{i \in I} \frac{\partial f}{\partial x_i}(a)\, dx_i$ был тождественно равен нулю.

Докажем теперь теорему в общем случае. Если f в точке a имеет относительный максимум или минимум, то функция $t \to f(a + t\vec{X})$ при фиксированном $\vec{X}$ в $\vec{E}$ определена для $t \in \mathbb{R}$ при достаточно малом $|t|$ (поскольку Ω — открытое множество) и имеет в точке $t = 0$ относительный максимум или минимум. Тогда доказательство, аналогичное доказательству теоремы Ролля (теорема 4), показывает, что ее производная в точке t равна нулю при $t = 0$. Эта производная равна $f'(a)\vec{X}$. Так как $f'(a)\vec{X} = 0$ для любого $\vec{X} \in \vec{E}$, то $f'(a) = 0$.

Замечания. По многим причинам теорема 23 не решает задачи об отыскании экстремума функции f.

1°) Чаще всего вещественная функция f задается не на открытом, а на замкнутом множестве F (например, на сегменте $[a, b]$ из $\mathbb{R}$ или в замкнутом шаре из E). Она может и не иметь никакого экстремума, если множество, на котором она определена, не компактно (например, $f(x) = x$ не имеет экстремума на $\mathbb{R}$). Если она и имеет экстремум, то он может не удовлетворять теореме, *которая применима только к открытым множествам.* Естественно, если f дифференцируема во внутренности $\Omega = \overset{\circ}{F}$ множества F, то каждая точка экстремума a, находящаяся в Ω, будет найдена из условия $f'(a) = 0$, но экстремальные значения могут достигаться на границе множества F (например, если f есть функция x^2, определенная на отрезке $[-1, 1]$, то ее минимум при $x = 0$ находится приравниванием нулю производной, поскольку 0 лежит внутри отрезка $[-1, 1]$, но точки максимума $x = \pm 1$ теореме не удовлетворяют, поскольку они находятся на концах этого интервала, где производная в нуль не обращается) [1].

2°) Обратно, уравнение $f'(a) = 0$ зачастую определяет точки, не являющиеся точками экстремума. Другими словами, необходимое условие $f'(a) = 0$ далеко не достаточно. Чтобы узнать, как ведет себя функция в точке a, надо изучить ее разложение Тейлора в окрестности этой точки.

---

[1]) В начале § 2 на стр. 195 мы говорили, что *большинство* теорем, сформулированных для открытого множества Ω из E, применимы также и к другим множествам, в особенности к неоткрытым интервалам $\mathbb{R}$. *Теорема* 23 — *одна из тех, которые в этих условиях неприменимы.* Ее доказательство опирается на доказательство теоремы Ролля (теорема 4), в котором существенно используется тот факт, что экстремальное значение достигается во внутренней точке интервала.

Исследуем подробнее эту проблему. Иногда факт обращения в нуль производной $f'(a)$ *более важен, чем свойство экстремальности.* Например, если рассматривается гиперповерхность $y=f(x)$ в пространстве $\Omega\times\mathbb{R}$, то условие $f'(a)=0$ означает, что касательная гиперплоскость, проведенная к поверхности в точке $A=(a,f(a))$, горизонтальна. Это необходимое условие экстремума ординаты геометрически более важно, чем само свойство экстремальности. Поэтому *независимо от того, имеется или нет в точке a экстремум функции, мы будем называть точку a стационарной точкой функции f, если в ней* $f'(a)=0$.

Нахождение необходимых и достаточных условий экстремума функции

Теорема 24. *Пусть f — вещественная функция, определенная на открытом множестве Ω аффинного нормированного пространства E и m раз дифференцируемая. Если производные функции f порядка $1, 2, \ldots, m-1$ равны нулю в точке $a\in\Omega$, а $f^{(m)}(a)\neq 0$ и если в точке a функция f имеет относительный максимум, то m четно и справедливо неравенство*

$$f^{(m)}(a)\vec{X}^m\leqslant 0 \quad \text{для любого} \quad \vec{X}\in\vec{E}. \tag{III, 7; 27}$$

Обратно, если f дифференцируема m раз в Ω, все производные f порядка $1, 2, \ldots, m-1$ равны нулю в точке $a\in\Omega$ и если $f^{(m)}(a)\vec{X}^m$ ограничено сверху некоторым числом $-\delta<0$ для всех $\vec{X}$ из единичной сферы $\|\vec{X}\|=1$ пространства $\vec{E}$, то f имеет в точке a строгий относительный максимум.

Доказательство. 1°) Докажем сначала первую часть теоремы, предполагая, что f имеет в точке $a\in\Omega$ относительный максимум. В этом случае функция $g\colon t\to f(a+t\vec{X})$, определенная в окрестности точки $t=0$ на вещественной прямой $\mathbb{R}$, имеет относительный максимум в точке $t=0$. Ее производная порядка k при $t=0$ имеет вид $f^{(k)}(a)\vec{X}^k$. При $t=0$ производные порядков $\leqslant m-1$ равны нулю, а производная порядка m равна $f^{(m)}(a)\vec{X}^m$. Поэтому для t, близких к нулю, имеем:

$$g(t)-g(0)=(f^{(m)}(a)\vec{X}^m)\frac{t^m}{m!}+\varepsilon_t t^m, \tag{III, 7; 28}$$

где ε_t стремится к 0 вместе с $|t|$. Отсюда следует, что либо $f^{(m)}(a)\vec{X}^m=0$, либо разность $g(t)-g(0)$ в окрестности точки $t=0$ имеет тот же знак, что и $(f^{(m)}(a)\vec{X}^m)t^m$. Поэтому, окончательно, при любом $\vec{X} \in \vec{E}$ и вещественном t имеет место неравенство

$$(f^{(m)}(a)\vec{X}^m)\, t^m \leqslant 0. \qquad \text{(III, 7; 29)}$$

При четном m это неравенство эквивалентно (III, 7; 27). Для доказательства первой части остается показать, что m не может быть нечетным. В самом деле, при нечетном m, полагая последовательно $t = +1$ и $t = -1$, из (III, 7; 29) получаем

$$f^{(m)}(a)\vec{X}^m = 0 \quad \text{для любого} \quad \vec{X} \in \vec{E}. \qquad \text{(III, 7; 30)}$$

Согласно лемме, относящейся к теореме 21_2, это соотношение влечет за собой равенство

$$f^{(m)}(a)(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_m) = 0 \quad \text{для любых} \quad (\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_m) \in \vec{E}^m, \qquad \text{(III, 7; 31)}$$

или $f^{(m)}(a) = 0$, что противоречит исходному предположению.

Замечание. Перед тем как перейти ко второй части теоремы, заметим следующее: может показаться, что если f имеет в точке *a строгий* относительный максимум, то форма порядка m: $f^{(m)}(a)\vec{X}^m$ *является отрицательно определенной*, т. е. имеет место неравенство

$$f^{(m)}(a)\vec{X}^m < 0 \quad \text{для} \quad \vec{X} \neq \vec{0}. \qquad \text{(III, 7; 35)}$$

Но это не так (за исключением, очевидно, случая, когда E одномерно). В самом деле, рассмотрим вещественную функцию двух переменных x, y:

$$f(x, y) = -(x^2 + y^4). \qquad \text{(III, 7; 36)}$$

В начале координат она имеет строгий максимум. Поскольку это полином, его разложение по степеням x и y является разложением Маклорена, так что, обозначая через $\overrightarrow{(X, Y)}$ вектор из $\vec{E} = \mathbb{R}^2$, получаем, что

$$\frac{1}{2!} f''(0, 0)\overrightarrow{(X, Y)}^2 = -X^2 \qquad \text{(III, 7; 37)}$$

всегда $\leqslant 0$, но не является отрицательно определенной (это выражение обращается в нуль на векторах $\overrightarrow{(0, Y)} \neq \overrightarrow{(0, 0)}$).

2°) Докажем теперь вторую часть теоремы. Заметим прежде всего, что *условия второй части более сильные, чем заключение первой.* Иначе не может и быть. Если предположить, что $f^{(k)}(a) = 0$ при $k = 1, 2, \ldots, m-1$, $f^{(m)}(a) \neq 0$ и что выполняется неравенство (III, 7; 27), то этого вовсе не достаточно для того, чтобы обеспечить функции f относительный максимум в точке a.

В самом деле, рассмотрим функцию f двух вещественных переменных x, y:

$$f(x, y) = -x^2 + y^4. \qquad (III, 7; 38)$$

Как мы видели для (III, 7; 36), имеет место соотношение (III, 7; 37), т. е. неравенство (III, 7; 27). Однако $f(0, 0) = 0$ и $f(0, y) > 0$ для $y \neq 0$, а, следовательно, f в начале координат относительного максимума не имеет. Поэтому мы ставим более сильные условия, указанные в формулировке теоремы: $f^{(m)}(a)\vec{X}^m$ мажорируется на единичной сфере числом $-\delta < 0$. Это влечет за собой, в силу однородности, неравенство

$$f^{(m)}(a)\vec{X}^m \leqslant -\delta \|\vec{X}\|^m. \qquad (III, 7; 39)$$

Последнее, в свою очередь, влечет за собой (III, 7; 35): $f^{(m)}(a)$ отрицательно определена. *Если E конечномерно, то неравенство* (III, 7; 39) *эквивалентно неравенству* (III, 7; 35). Действительно, в этом случае единичная сфера компактна и непрерывная функция $\vec{X} \to f^{(m)}(a)\vec{X}^m$, отрицательная в силу (III, 7; 35) на этой сфере, достигает на ней максимума, являющегося некоторым числом $-\delta < 0$. Однако для бесконечномерного пространства E это условие более сильное, чем (III, 7; 35).

Запишем разложение Тейлора порядка m в виде (III, 7; 10):

$$f(a + \vec{h}) = f(a) + \frac{f^{(m)}(a)}{m!}\vec{h}^m + \alpha(\vec{h})\|\vec{h}\|^m, \qquad (III, 7; 40)$$

где $\alpha(\vec{h})$ стремится к $\vec{0}$ вместе с $\vec{h}$. Тогда (III, 7; 39) дает

$$f(a + \vec{h}) \leqslant f(a) - \delta\|\vec{h}\|^m + \alpha(\vec{h})\|\vec{h}\|^m. \qquad (III, 7; 41)$$

Так как $\vec{\alpha}$ стремится к $\vec{0}$, то для $\vec{h} \neq \vec{0}$ с достаточно малой нормой $f(a + \vec{h}) - f(a) < 0$ и, следовательно, в точке a функция f имеет строгий относительный максимум.

Естественно, аналогичная теорема имеет место для минимума. Отсюда получается правило, позволяющее выяснить, имеет ли f в точке a относительный максимум или минимум.

Правило. *Прежде всего необходимо, чтобы $f'(a) = 0$. Затем ищется первое целое число m, такое, что $f^{(m)}(a) \neq 0$. Если m нечетно, то в точке a нет ни относительного максимума, ни относительного минимума. Если m четно, то рассматривается знак формы степени m: $\vec{X} \to f^{(m)}(a)\vec{X}^m$. Если она может принимать значения разных знаков, то в точке a нет ни относительного* *максимума, ни относительного минимума*[1]). *Если она всегда* $\leqslant 0$ (*соответственно* $\geqslant 0$), *не будучи отрицательно определенной* (*соответственно положительно определенной*), *то без изучения производных порядка* $>m$ *ничего определенного заключить нельзя. Если же она отрицательно определенная* (*соответственно* >0) *и* E *конечномерно или же если имеет место более сильное условие* $\sup\limits_{\|\vec{X}\|=1} (f^{(m)}(a)\vec{X}^m)<0$ (*соответственно* $\inf\limits_{\|\vec{X}\|=1} (f^{(m)}(a)\vec{X}^m)>0$) *для бесконечномерного* E, *то* a *является точкой строгого относительного максимума* (*соответственно минимума*).

Из всего изложенного следует, что не существует простого условия, одновременно необходимого и достаточного для того, чтобы вещественная функция, определенная на $\Omega \subset E$, имела в точке a относительный максимум или минимум.

Замечания. 1°) Если функция f *имеет в точке* a *все последовательные производные, равные нулю*, и не является тождественно нулевой функцией, то изучение разложения Тейлора не позволяет выяснить вопроса о существовании в точке a относительного максимума или относительного минимума.

2°) Предположим, что в точке a функция имеет седло. Тогда можно попытаться узнать, в какой части области Ω в окрестности точки a имеет место неравенство $f(x) \geqslant f(a)$ и в какой части $f(x) \leqslant f(a)$. Здесь применимы те же правила. *Если на множестве* B *единичной сферы* $\|\vec{X}\|=1$ *величина* $f^{(m)}(a)\vec{X}^m$ *не превосходит* $-\delta<0$, *то во всем конусе* $\{a+\lambda\vec{X};\ \vec{X}\in B,\ \lambda>0$ *вещественно*$\}$ *в окрестности точки* a *имеет место неравенство* $f(x)<f(a)$.

Частный случай вещественной функции f двух вещественных переменных x, y

При исследовании такой функции на экстремум надо прежде всего найти точки, в которых частные производные $p=\frac{\partial f}{\partial x}$ и $q=\frac{\partial f}{\partial y}$ равны нулю. Пусть (a, b) — такая точка. Рассмот-

[1]) В этом случае говорят, что в точке a функция f имеет *седло*. Типичный пример такого седла дает функция двух вещественных переменных $f(x, y)=xy$. Ее графиком в $\mathbb{R}^3$ служит гиперболический параболоид $z=xy$. В начале координат функция f имеет седло. Сужением f на прямую $y=kx$ является функция $x\to kx^2$, имеющая строгий максимум в начале координат при $k<0$ и строгий минимум при $k>0$. Можно также говорить о седле и в том случае, когда m нечетно.

рим теперь разложение функции f по формуле Тейлора по степеням $x-a=X$ и $y-b=Y$. Это разложение начинается с

$$\frac{1}{2}(rX^2+2sXY+tY^2). \tag{III,7;42}$$

1-*й случай*: $rt-s^2>0$, r (и t) <0. Квадратичная форма (III, 7; 42) в этом случае отрицательно определенная, а, значит, f имеет в точке (a, b) строгий относительный максимум.

2-*й случай*: $rt-s^2>0$, r (и t) >0. Функция f имеет в точке (a, b) строгий относительный минимум.

3-*й случай*: $rt-s^2<0$. Квадратичная форма (III, 7; 42) принимает значения разных знаков, и, следовательно, мы имеем дело с *седлом*. Уравнение $rX^2+2sXY+tY^2=0$ задает две прямые D и D' на плоскости $\mathbb{R}^2$, определяющие 4 попарно вертикальных угла. Обозначим их по порядку обхода через (1), (2), (3), (4). Предположим, например, что рассматриваемая квадратичная форма $\geqslant 0$ в углах (1) и (3) и $\leqslant 0$ в углах (2) и (4).

Пусть угол (1) определен в полярных координатах неравенствами $\rho\geqslant 0$ и $\alpha\leqslant\varphi\leqslant\beta$. Тогда каждый угол $\alpha'\leqslant\varphi\leqslant\beta'$, где $\alpha'>\alpha$ и $\beta'<\beta$, пересекает «единичную сферу» из $\mathbb{R}^2$, т. е. окружность, по *компактной* дуге, на которой квадратичная форма имеет минимум $\delta>0$. Следовательно, существует такое число $\varepsilon>0$, что если положить $x=a+\rho\cos\varphi$, $y=b+\rho\sin\varphi$, $\rho>0$, $\alpha'\leqslant\varphi\leqslant\beta'$, то для $0<\rho\leqslant\varepsilon$ имеет место неравенство $f(x,y)>f(a,b)$.

Однако *это число* ε *зависит от* α' *и* β'. Поэтому невозможно, не привлекая к изучению производные f порядков >2 в точке (a, b), делать какие-либо выводы при ρ, стремящемся к 0, если одновременно с этим φ стремится к α или β. Тот же результат имеет место в угле (3). В углах (2) и (4) имеет место аналогичный результат, но с заменой неравенства $f(x, y)>f(a,b)$ на $f(x,y)<f(a,b)$.

Области плоскости $\mathbb{R}^2$, в которых $f(x,y)>f(a,b)$ и $f(x,y)<f(a,b)$, разделены кривой $f(x,y)=f(a,b)$. Эта кривая в точке (a,b) имеет две ветви, касательные соответственно к D и D'. Однако, не привлекая к рассмотрению производные функции f порядков >2 в точке (a,b), невозможно сказать, в каком из углов находятся эти кривые.

4-*й случай*: $rt-s^2=0$, r или $t<0$. Квадратичная форма равна квадрату некоторой линейной формы $\neq 0$, взятому с противоположным знаком. Она $\leqslant 0$, но обращается в нуль на некоторой прямой D. Пусть $\varphi=\varphi_0+k\pi$ — полярный угол этой прямой. Если $x=a+\rho\cos\varphi$, $y=b+\rho\sin\varphi$, $\rho\geqslant 0$, $\alpha'\leqslant\varphi\leqslant\beta'$, где α' и β' фиксированы, $\alpha'>\varphi_0$ и $\beta'<\varphi_0+\pi$, то существует такое число $\varepsilon>0$, что для $0<\rho\leqslant\varepsilon$ имеет место неравенство

$f(x, y) < f(a, b)$. Аналогичный результат справедлив для $\alpha' > \varphi_0 + \pi$, $\beta' < \varphi_0 + 2\pi$.

К сказанному ничего добавить нельзя. В частности, без привлечения производных порядков >2 нельзя сказать, имеет или нет функция f в точке (a, b) относительный максимум.

5-*й случай*: $rt - s^2 = 0$, r или $t > 0$. Результат аналогичен предыдущему, надо только заменить $f(x, y) < f(a, b)$ на $f(x, y) > f(a, b)$.

6-*й случай*: $r = s = t = 0$. Без привлечения производных функции f в точке (a, b) порядка >2 никакого вывода о поведении функции f в точке (a, b) сделать невозможно.

Заключение. Проведенные исследования показывают, что различных вариантов поведения функции более одной переменной *значительно больше*, чем это имеет место для функции одной переменной.

Применение формулы Тейлора к изучению расположения гиперповерхности по отношению к касательной гиперплоскости

Пусть f — дифференцируемая вещественная функция, определенная на $\Omega \subset E$. Тогда $y = f(x)$ является уравнением дифференцируемого многообразия пространства $E \times \mathbb{R}$. Такое многообразие называют *дифференцируемой гиперповерхностью*. Ее касательная гиперплоскость в точке $A = (a, f(a))$ определяется уравнением (III, 3; 19_3). Расположение гиперповерхности по отношению к касательной гиперплоскости в окрестности точки A определяется знаком разности

$$\overrightarrow{f(x) - f(a)} - f'(a)\,\overrightarrow{(x - a)} \qquad \text{(III, 7; 43)}$$

для x, близких к a.

Но $x \to f(x) - f'(a)\,\overrightarrow{(x - a)}$ является дифференцируемой вещественной функцией, производная которой в точке a равна нулю. Все сводится к тому, чтобы выяснить, имеет ли эта функция в точке a экстремум какого-нибудь типа или седло. Тем самым *мы возвращаемся к теореме* 24. В силу приведенного выше заключения, наличие седла является обстоятельством частым, а не исключительным и, следовательно, кроме случая одномерного E, гиперповерхность обычно пересекает свою касательную гиперплоскость, даже если $f''(a) \neq 0$. В частном случае функции f двух переменных x, y поверхность локально расположена с одной стороны от касательной плоскости, если $rt - s^2 > 0$, и пересекает ее, если $rt - s^2 < 0$.