В этом случае элементы матрицы A_1 однозначно определятся из уравнений (92):

$$x_{ik}^{(1)} = -\frac{p_{ik}^{(0)}}{\lambda_i - \lambda_k - 1} \quad (i,\ k = 1,\ 2,\ \ldots,\ n). \tag{93}$$

Если же при некоторых i, k[1])

$$\lambda_i - \lambda_k = 1,$$

то соответствующее $p_{ik}^{(0*)}$ определяется из (92):

$$p_{ik}^{(0*)} = p_{ik}^{(0)},$$

а соответствующее $x_{ik}^{(1)}$ выбирается совершенно произвольно.

При тех же i и k, при которых $\lambda_i - \lambda_k \neq 1$, мы полагаем:

$$p_{ik}^{(0*)} = 0,$$

а соответствующее $x_{ik}^{(1)}$ находим по формуле (93).

Определив A_1, мы переходим к определению A_2 из третьего уравнения (87). Заменим это матричное уравнение системой n^2 скалярных уравнений:

$$(\lambda_i - \lambda_k - 2)\, x_{ik}^{(2)} = p_{ik}^{(1*)} - p_{ik}^{(1)} - (P_0 A_1 - A_1 P_0^*)_{ik} \tag{94}$$
$$(i,\ k = 1,\ 2,\ \ldots,\ n).$$

Здесь мы поступаем так же, как и при определении A_1.

Если $\lambda_i - \lambda_k \neq 2$, то мы полагаем:

$$p_{ik}^{(1*)} = 0,$$

и тогда из (94) находим:

$$x_{ik}^{(2)} = -\frac{1}{\lambda_1 - \lambda_2 - 2}\,[p_{ik}^{(1)} - (P_0 A_1 - A_1 P_0^*)_{ik}].$$

Если же $\lambda_i - \lambda_k = 2$, то при этих i и k из (94) следует, что

$$p_{ik}^{(1*)} = p_{ik}^{(1)} + (P_0 A_1 - A_1 P_0^*)_{ik}.$$

В этом случае $x_{ik}^{(2)}$ выбирается произвольно.

Продолжая этот процесс далее, мы последовательно определим все матрицы P_{-1}^*, P^*, P_1^*, ... и A_1, A_2, ...

При этом только конечное число из матриц P_m^* будет отлично от нуля, и, как нетрудно видеть, матрица $P^*(z)$ будет иметь вид[2])

$$P^*(z) = \begin{Vmatrix} \frac{\lambda_1}{z} & a_{12} z^{\lambda_1 - \lambda_2 - 1} & \ldots & a_{1n} z^{\lambda_1 - \lambda_n - 1} \\ 0 & \frac{\lambda_2}{z} & \ldots & a_{2n} z^{\lambda_2 - \lambda_n - 1} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \frac{\lambda_n}{z} \end{Vmatrix}, \tag{95}$$

[1]) В силу (89) это возможно лишь при $i < k$.

[2]) P_m^* $(m \geqslant 0)$ может быть отличным от нуля лишь тогда, когда существуют характеристические числа λ_i и λ_k матрицы P_{-1} такие, что $\lambda_i - \lambda_k - 1 = m$ (при этом в силу (89) $i < k$). При данном m каждому такому равенству соответствует элемент $p_{ik}^{(m*)} = a_{ik}$ матрицы P_m^*; этот элемент может быть отличен от нуля. Все остальные элементы матрицы P_m^* равны нулю.

где $a_{ik}=0$, если $\lambda_i-\lambda_k$ не есть целое положительное число, и $a_{ik}=p_{ik}^{(\lambda_i-\lambda_k-1^*)}$, если $\lambda_i-\lambda_k$ является целым положительным числом.

Обозначим через m_i наибольшую целую часть числа $\operatorname{Re}\lambda_i$[1]):

$$m_i=[\operatorname{Re}\lambda_i] \qquad (i=1,\ 2,\ \ldots,\ n). \tag{96}$$

Тогда в силу (89)

$$m_1\geqslant m_2\geqslant\ldots\geqslant m_n.$$

При этом если $\lambda_i-\lambda_k$ есть целое число, то

$$\lambda_i-\lambda_k=m_i-m_k.$$

Поэтому в выражении (95) канонической матрицы $P^*(z)$ мы можем все разности $\lambda_i-\lambda_k$ заменить на m_i-m_k. Кроме того, мы положим:

$$\tilde{\lambda}_i=\lambda_i-m_i \qquad (i=1,\ 2,\ \ldots,\ m), \tag{91'}$$

$$M=\|m_i\delta_{ik}\|_1^n, \qquad U=\begin{Vmatrix}\tilde{\lambda}_1 & a_{12} & \ldots & a_{1n}\\ 0 & \tilde{\lambda}_2 & \ldots & a_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & \tilde{\lambda}_n\end{Vmatrix}. \tag{97}$$

Тогда из (95) следует (см. формулу I на стр. 435):

$$P^*(z)=z^M\frac{U}{z}z^{-M}+\frac{M}{z}=D_z(z^Mz^U).$$

Отсюда вытекает, что $Y=z^Mz^U$ представляет собой решение уравнения (85), а

$$X=A(z)\,z^Mz^U \tag{98}$$

является решением уравнения (81)[2]).

3° Переходим к общему случаю. Как было выяснено выше, мы можем, не нарушая общности, матрицу P_{-1} заменить любой матрицей, ей подобной. Мы примем, что матрица P_{-1} имеет нормальную жорданову форму[3])

$$P_{-1}=\{\lambda_1E_1+H_1,\ \lambda_2E_2+H_2,\ \ldots,\ \lambda_uE_u+H_u\}, \tag{99}$$

причем

$$\operatorname{Re}\lambda_1\geqslant\operatorname{Re}\lambda_2\geqslant\ldots\geqslant\operatorname{Re}\lambda_u. \tag{100}$$

Здесь E обозначает единичную матрицу, а H — матрицу, у которой элементы первой «наддиагонали» равны единице, а остальные элементы равны нулю. Порядки матриц E_i и H_i в различных диагональных клетках будут, вообще говоря, различными; эти порядки совпадают со степенями соответствующих элементарных делителей матрицы P_{-1}[4]).

В соответствии с представлением (99) матрицы P_{-1} разобьем все матрицы P_m, P_m^*, A_m на блоки:

$$P_m=(P_{ik}^{(m)})_1^u,\ \ P_m^*=(P_{ik}^{(m^*)})_1^u,\ \ A_m=(X_{ik}^{(m)})_1^u.$$

---

[1]) То есть m_i — наибольшее целое число, не превосходящее $\operatorname{Re}\lambda_i$ $(i=1,\ 2,\ \ldots,\ n)$.

[2]) Специальный вид матриц (97) соответствует каноническому виду матрицы P_{-1}. Если матрица P_{-1} не имеет каноническую форму, то матрицы M и U в (98) подобны матрицам (97).

[3]) См. гл. VI, § 6.

[4]) Для сокращения обозначений мы не пишем при E_i и H_i индекса, указывающего порядок этих матриц.

Тогда второе из уравнений (87) может быть заменено системой уравнений

$$(\lambda_i E_i + H_i) X_{ik}^{(1)} - X_{ik}^{(1)} [(\lambda_k + 1) E_k + H_k] + P_{ik}^{(0)} = P_{ik}^{(0*)} \tag{101}$$
$$(i, k = 1, 2, \ldots, u),$$

которые могут быть еще переписаны так:

$$(\lambda_i - \lambda_k - 1) X_{ik}^{(1)} + H_i X_{ik}^{(1)} - X_{ik}^{(1)} H_k + P_{ik}^{(0)} = P_{ik}^{(0*)} \quad (i, k = 1, 2, \ldots, u). \tag{102}$$

Пусть [1])

$$X_{ik}^{(1)} = \begin{Vmatrix} x_{11} & x_{12} & \ldots \\ x_{21} & x_{22} & \ldots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{Vmatrix} = \| x_{st} \|, \quad P_{ik}^{(0)} = \| p_{st}^{(0)} \|, \quad P_{ik}^{(0*)} = \| p_{st}^{(0*)} \|.$$

Тогда матричное уравнение (102) (при фиксированных i и k) можно заменить системой скалярных уравнений вида [2])

$$(\lambda_i - \lambda_k - 1) x_{st} + x_{s+1,\, t} - x_{s,\, t-1} + p_{st}^{(0)} = p_{st}^{(0*)} \tag{103}$$
$$(x_{v+1,\, t} = x_{s,\, 0} = 0; \quad s = 1, 2, \ldots, v; \quad t = 1, 2, \ldots, w),$$

где v и w — порядки матриц $\lambda_i E_i + H_i$ и $\lambda_k E_k + H_k$ в (99).

Если $\lambda_i - \lambda_k \neq 1$, то в системе (103) можно положить все $p_{st}^{(0*)} = 0$ и однозначно определить все x_{st} из рекуррентных соотношений (103). Это означает, что в матричном уравнении (102) мы полагаем

$$P_{ik}^{(0*)} = 0$$

и однозначно определяем $X_{ik}^{(1)}$.

Если $\lambda_i - \lambda_k = 1$, то соотношения (103) принимают вид

$$x_{s+1,\, t} - x_{s,\, t-1} + p_{st}^{(0)} = p_{st}^{(0*)} \tag{104}$$
$$(s = 1, 2, \ldots, v; \; t = 1, 2, \ldots, w; \; x_{v+1,\, t} = x_{s,\, 0} = 0).$$

Нетрудно показать, что из уравнений (104) можно так определить элементы x_{st} матрицы $X_{ik}^{(1)}$, чтобы матрица $P_{ik}^{(0*)}$ имела в соответствии со своими размерами $(v \times w)$ вид

$$\begin{Vmatrix} a_0 & 0 & \ldots & 0 \\ a_1 & a_0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{v-1} & a_{v-2} & \ldots & a_1 \;\; a_0 \end{Vmatrix}, \qquad \begin{Vmatrix} a_0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ a_1 & a_0 & \ldots & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{v-1} & & & a_1 & a_0 & 0 & \ldots 0 \end{Vmatrix},$$

$$(v = w) \qquad\qquad (v < w)$$

[1]) Для сокращения обозначений мы опускаем индексы i, k при обозначении элементов матриц X_{ik}, $P_{ik}^{(0)}$, $P_{ik}^{(0*)}$.

[2]) Рекомендуем читателю вспомнить свойства матрицы H, разобранные на стр. 24—25.

$$\begin{Vmatrix} 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 0 \\ a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{w-1} & \dots & a_1 & a_0 \end{Vmatrix}. \tag{105}$$

$$(v > w)$$

Про матрицы (105) будем говорить, что они имеют *правильную нижнюю треугольную форму*[1]).

Из третьего уравнения (87) мы определяем матрицу A_2. Это уравнение можно заменить системой уравнений

$$(\lambda_i - \lambda_k - 2)\, X_{ik}^{(2)} + H_i X_{ik}^{(2)} - X_{ik}^{(2)} H_k + \{P_0 A_1 - A_1 P_0\}_{ik} + P_{ik}^{(1)} = P_{ik}^{(1*)} \tag{106}$$

$$(i,\ k = 1,\ 2,\ \dots,\ u).$$

Аналогично тому, как это было при определении A_1, если $\lambda_i - \lambda_k \neq 2$, то из соответствующего уравнения (106) матрица $X_{ik}^{(2)}$ определяется однозначно при $P_{ik}^{(1*)} = 0$. Если же $\lambda_i - \lambda_k = 2$, то можно так определить матрицу $X_{ik}^{(2)}$, чтобы матрица $P_{ik}^{(1*)}$ имела правильную нижнюю треугольную форму.

Продолжая этот процесс, мы определим последовательно все матричные коэффициенты A_1, A_2, ... и P_{-1}^*, P_0^*, P_1^*, ... При этом только конечное число из коэффициентов P_m^* будет отлично от нуля, и матрица $P^*(z)$ будет иметь следующий блочный вид[2]):

$$P^*(z) = \begin{pmatrix} \dfrac{\lambda_1 E_1 + H_1}{z} & B_{12} z^{\lambda_1 - \lambda_2 - 1} & \dots & B_{1u} z^{\lambda_1 - \lambda_u - 1} \\ 0 & \dfrac{\lambda_2 E_2 + H_2}{z} & \dots & B_{2u} z^{\lambda_2 - \lambda_u - 1} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dfrac{\lambda_u E_u + H_u}{z} \end{pmatrix}, \tag{107}$$

где

$$B_{ik} = \begin{cases} 0, & \text{если } \lambda_i - \lambda_k \text{ не есть целое положительное число;} \\ P_{ik}^{(\lambda_i - \lambda_k - 1*)}, & \text{если } \lambda_i - \lambda_k = \text{целому положительному числу} \\ & (i,\ k = 1,\ 2,\ \dots,\ u). \end{cases}$$

---

[1]) Аналогично определяются верхние правильные треугольные матрицы.

Из уравнений (104) все элементы матрицы $X_{ik}^{(1)}$ не определяются однозначно; имеется некоторый произвол в выборе элементов x_{st}. Это видно и непосредственно из уравнения (102): при $\lambda_i - \lambda_k = 1$ к матрице $X_{ik}^{(1)}$ можно прибавить произвольную матрицу, перестановочную с H, т. е. произвольную правильную верхнюю треугольную матрицу.

[2]) Размеры квадратных матриц E_i, H_i и прямоугольных матриц B_{ik} определяются размерами диагональных клеток в жордановой матрице P_{-1}, т. е. степенями элементарных делителей матрицы P_{-1}.

Все матрицы B_{ik} $(i, k=1, 2, \ldots, u;\ i<k)$ имеют правильную нижнюю треугольную форму.

Как и в предыдущем случае, обозначим через m_i целую часть $\operatorname{Re}\lambda_i$

$$m_i=[\operatorname{Re}\lambda_i] \qquad (i=1,\ 2,\ \ldots,\ u) \tag{108}$$

и положим

$$\lambda_i=m_i+\tilde{\lambda}_i \qquad (i=1,\ 2,\ \ldots,\ u). \tag{108'}$$

Тогда снова в выражении (107) для $P^*(z)$ мы можем всюду разность $\lambda_i-\lambda_k$ заменить разностью m_i-m_k. Вводя диагональную матрицу с целыми элементами M и верхнюю треугольную матрицу U при помощи равенств[1])

$$M=(m_iE_i\delta_{ik})_1^u, \quad U=\begin{pmatrix} \tilde{\lambda}_1E_1+H_1 & B_{12} & \ldots & B_{1u} \\ 0 & \tilde{\lambda}_2E_2+H_2 & \ldots & B_{2u} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \tilde{\lambda}_uE_u+H_u \end{pmatrix}, \tag{109}$$

мы, исходя из (107), легко получим следующее представление для матрицы $P^*(z)$:

$$P^*(z)=z^M\frac{U}{z}\cdot z^{-M}+\frac{M}{z}=D_z(z^Mz^U).$$

Отсюда следует, что решение уравнения (85) может быть задано в виде

$$Y=z^Mz^U,$$

а решение уравнения (81) может быть представлено так:

$$X=A(z)\,z^Mz^U. \tag{110}$$

Здесь $A(z)$ — матричный ряд (84), M — диагональная матрица с постоянными целыми элементами, U — постоянная треугольная матрица. Матрицы M и U определяются равенствами (108), (108′) и (109)[2]).

3. Переходим теперь к доказательству сходимости ряда

$$A(z)=E+A_1z+A_2z^2+\ldots$$

Воспользуемся леммой, которая представляет и самостоятельный интерес.

Лемма. *Если ряд*

$$x=a_0+a_1z+a_2z^2+\ldots \tag{111}$$

формально удовлетворяет системе[3])

$$\frac{dx}{dz}=P(z)\,x, \tag{112}$$

для которой $z=0$ является регулярной особой точкой, то ряд (111) *сходится в любой окрестности точки $z=0$, в которой сходится разложение в ряд* (82) *для матрицы коэффициентов $P(z)$.*

---

[1]) Здесь разбиение на блоки соответствует разбиению матриц P_{-1} и $P^*(z)$.

[2]) См. сноску [2]) к стр. 452.

[3]) Здесь $x=(x_1, x_2, \ldots, x_n)$ — столбец из неизвестных функций; $a_0, a_1, a_2, \ldots$ — постоянные столбцы; $P(z)$ — квадратная матрица коэффициентов.

Доказательство. Пусть

$$P(z)=\frac{P_{-1}}{z}+\sum_{q=0}^{\infty} P_q z^q,$$

где ряд $\sum\limits_{m=0}^{\infty} P_m z^m$ сходится при $|z|<r$. Тогда существуют такие положительные постоянные p_{-1} и p, что [1])

$$\operatorname{mod} P_{-1}\leqslant p_{-1} I,\ \operatorname{mod} P_m\leqslant \frac{p}{r^m} I,\ I=\|1\| \quad (m=0,\ 1,\ 2,\ \ldots). \tag{113}$$

Подставляя в (112) вместо x ряд (111) и приравнивая между собой коэффициенты при одинаковых степенях в обеих частях равенства (112), получим бесконечную систему векторных (столбцевых) равенств

$$\left.\begin{aligned} P_{-1}a_0&=0,\\ (E-P_{-1})\,a_1&=P_0a_0,\\ (2E-P_{-1})\,a_2&=P_0a_1+P_1a_0,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ (mE-P_{-1})\,a_m&=P_0a_{m-1}+P_1a_{m-2}+\ldots+P_{m-1}a_0\\ &\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}\right\} \tag{114}$$

Нам достаточно доказать, что какой-либо остаток ряда (111)

$$x^{(k)}=a_kz^k+a_{k+1}z^{k+1}+\ldots \tag{115}$$

сходится в окрестности точки $z=0$. Число k подчиним неравенству

$$k>np_{-1}.$$

Тогда число k будет превосходить модули всех характеристических чисел матрицы P_{-1} [2]) и потому при $m\geqslant k$ будем иметь $|mE-P_{-1}|\neq 0$ и

$$(mE-P_{-1})^{-1}=\frac{1}{m}\left(E-\frac{P_{-1}}{m}\right)^{-1}=\frac{1}{m}E+\frac{1}{m^2}P_{-1}+\frac{1}{m^3}P_{-1}^2+\ldots \tag{116}$$

$$(m=k,\ k+1,\ \ldots).$$

В последней части этого равенства стоит сходящийся матричный ряд. Пользуясь этим рядом, мы из (114) можем все коэффициенты ряда (115) выразить однозначно через $a_0,\ a_1,\ \ldots,\ a_{k-1}$ при помощи рекуррентных

---

[1]) Относительно определения модуля матрицы см. стр. 431.

[2]) Если λ_0 — характеристическое число матрицы $A=\|a_{ik}\|_1^n$, то $|\lambda_0|\leqslant n\max\limits_{1\leqslant i,\ k\leqslant n}|a_{ik}|$. Действительно, пусть $Ax=\lambda_0x$, где $x=(x_1,\ x_2,\ \ldots,\ x_n)\neq 0$. Тогда

$$\lambda_0x_i=\sum_{k=1}^{n} a_{ik}x_k \qquad (i=1,\ 2,\ \ldots,\ n).$$

Пусть $|x_j|=\max\{|x_1|,\ |x_2|,\ \ldots,\ |x_n|\}$. Тогда

$$|\lambda_0|\,|x_j|\leqslant\sum_{k=1}^{n}|a_{jk}|\,|x_k|\leqslant|x_j|\,n\max_{1\leqslant i,\ k\leqslant n}|a_{ik}|.$$

Сокращая на $|x_j|$, получим нужное неравенство.

соотношений

$$a_m = \left(\frac{1}{m}E + \frac{1}{m^2}P_{-1} + \frac{1}{m^3}P_{-1}^2 + \ldots\right)(f_{m-1} + P_0 a_{m-1} + \ldots + P_{m-k-1}a_k)$$
$$(m = k,\ k+1,\ \ldots), \tag{117}$$

где

$$f_{m-1} = P_{m-k}a_{k-1} + \ldots + P_{m-1}a_0 \quad (m = k,\ k+1,\ \ldots). \tag{118}$$

Заметим, что ряд (115) формально удовлетворяет дифференциальному уравнению

$$\frac{dx^{(k)}}{dz} = P(z)\,x^{(k)} + f(z), \tag{119}$$

где

$$f(z) = \sum_{m=k-1}^{\infty} f_m z^m = P(z)(a_0 + a_1 z + \ldots + a_{k-1}z^{k-1}) -$$
$$- a_1 - 2a_2 z - \ldots - (k-1)\,a_{k-1}z^{k-2}. \tag{120}$$

Из (120) вытекает, что ряд

$$\sum_{m=k-1}^{\infty} f_m z^m$$

сходится при $|z| < r$, и потому существует такое число $N > 0$, что[1]

$$\operatorname{mod} f_m \leqslant \left\| \frac{N}{r^m} \right\| \quad (m = k-1,\ k,\ \ldots). \tag{121}$$

Из вида рекуррентных соотношений (117) следует, что, заменив в них матрицы P_{-1}, P_q, f_{m-1} мажорантными матрицами $p_{-1}I$, $\frac{p}{r^q}I$, $\left\| \frac{N}{r^{m-1}} \right\|$, а столбец a_m столбцом $\|\alpha_m\|$[2] $(m = k,\ k+1,\ \ldots;\ q = 0,\ 1,\ 2,\ \ldots)$, мы получим соотношения, определяющие верхние границы $\|\alpha_m\|$ для $\operatorname{mod} a_m$:

$$\operatorname{mod} a_m \leqslant \|\alpha_m\|. \tag{122}$$

Следовательно, ряд

$$\xi^{(k)} = \alpha_k z^k + a_{k+1}z^{k+1} + \ldots \tag{123}$$

после почленного умножения на столбец $\|1\|$ будет мажорантным рядом для ряда (115).

Заменив в (119) матричные коэффициенты P_{-1}, P_q, f_m рядов

$$P(z) = \frac{P_{-1}}{z} + \sum_{q=0}^{\infty} P_q z^q, \qquad f(z) = \sum_{m=k-1}^{\infty} f_m z^m$$

соответствующими мажорантными матрицами $p_{-1}I$, $\frac{p}{r^q}I$, $\left\| \frac{N}{r^m} \right\|$, а также

---

[1]) Здесь $\left\| \frac{N}{r^m} \right\|$ обозначает столбец, у которого все элементы равны одному и тому же числу $\frac{N}{r^m}$.

[2]) Здесь $\|\alpha_m\|$ обозначает столбец $(\alpha_m, \alpha_m, \ldots, \alpha_m)$ (α_m — число; $m = k,\ k+1,\ \ldots$)

заменив $x^{(k)}$ на $\|\xi^{(k)}\|$, мы получим дифференциальное уравнение для $\xi^{(k)}$:

$$\frac{d\xi^{(k)}}{dz}=n\left(\frac{p_{-1}}{z}+\frac{p}{1-\frac{z}{r}}\right)\xi^{(k)}+\frac{N\frac{z^{k-1}}{r^{k-1}}}{1-\frac{z}{r}}. \tag{124}$$

Это линейное дифференциальное уравнение имеет частное решение

$$\xi^{(k)}=\frac{N}{r^{k-1}}\frac{z^{np-1}}{\left(1-\frac{z}{r}\right)^{npr}}\int_0^z z^{k-np-1}\left(1-\frac{z}{r}\right)^{npr-1}dz, \tag{125}$$

которое регулярно в точке $z=0$ и в окрестности этой точки разлагается в сходящийся при $|z|<r$ степенной ряд (123).

Из сходимости мажорантного ряда (123) следует сходимость ряда (115) при $|z|<r$. Лемма доказана.

Замечание 1. Приведенное доказательство позволяет определить все регулярные в особой точке решения дифференциальной системы (112), если таковые существуют.

Для существования регулярных решений (не равных тождественно нулю) необходимо и достаточно, чтобы матрица вычетов P_{-1} имела целое неотрицательное характеристическое число. Если s — наибольшее такое целое характеристическое число, то из первых $s+1$ уравнений (114) можно определить не обращающиеся одновременно в нуль столбцы a_0, $a_1, \ldots, a_s$, поскольку определитель соответствующей системы линейных однородных уравнений равен нулю:

$$\Delta=|P_{-1}|\,|E-P_{-1}|\,|\ldots|\,sE-P_{-1}|=0.$$

Из остальных уравнений (114) столбцы a_{s+1}, $a_{s+2}, \ldots$ однозначно выразятся через $a_0, a_1, \ldots, a_s$. Полученный ряд (111) сходится согласно лемме. Таким образом, линейно независимые решения первых $s+1$ уравнений (114) определяют все линейно независимые регулярные в особой точке $z=0$ решения системы (112).

Если $z=0$ есть особая точка, то задание начального значения a_0 для регулярного в этой точке решения (111) (если таковое существует) не определяет однозначно этого решения. Однако решение, регулярное в регулярной особой точке, определяется однозначно, если заданы a_0, $a_1, \ldots, a_s$, т. е. если заданы начальные значения при $z=0$ самого решения и его первых s производных (s — наибольшее неотрицательное целое характеристическое число матрицы вычетов P_{-1}).

Замечание 2. Доказанная лемма сохраняет свою силу и при $P_{-1}=0$. В этом случае в доказательстве леммы в качестве p_{-1} можно взять любое положительное число. При $P_{-1}=0$ лемма утверждает известное положение о существовании регулярного решения в окрестности регулярной точки системы. В этом случае решение однозначно определяется заданием a_0.

4. Пусть дана система

$$\frac{dX}{dz}=P(z)X, \tag{126}$$

где

$$P(z)=\frac{P_{-1}}{z}+\sum_{m=0}^{\infty}P_m z^m$$

и ряд, стоящий в правой части, сходится при $|z|<r$.

Пусть, далее, полагая

$$X = A(z)Y \tag{127}$$

и подставляя вместо $A(z)$ ряд

$$A(z) = A_0 + A_1 z + A_2 z^2 + \ldots, \tag{128}$$

мы после формальных преобразований получаем:

$$\frac{dY}{dz} = P^*(z)Y, \tag{129}$$

где

$$P^*(z) = \frac{P^*_{-1}}{z} + \sum_{m=0}^{\infty} P^*_m z^m,$$

причем здесь, как и в выражении для $P(z)$, ряд в правой части сходится при $|z| < r$.

Докажем, что и ряд (128) сходится в окрестности $|z| < r$ точки $z = 0$.

Действительно, из (126), (127) и (129) следует, что ряд (128) формально удовлетворяет следующему дифференциальному матричному уравнению

$$\frac{dA}{dz} = P(z)A - AP^*(z). \tag{130}$$

Мы будем рассматривать A как вектор (столбец) в пространстве всех матриц n-го порядка, т. е. в пространстве n^2 измерений. Если мы в этом пространстве определим линейный оператор $\widehat{P}(z)$ над матрицей A, аналитически зависящий от параметра z, при помощи равенства

$$\widehat{P}(z)[A] = P(z)A - AP^*(z), \tag{131}$$

то дифференциальное уравнение (130) можно будет записать в виде

$$\frac{dA}{dz} = \widehat{P}(z)[A]. \tag{132}$$

Правую часть этого уравнения можно рассматривать как произведение матрицы $\widehat{P}(z)$ порядка n^2 на столбец A из n^2 элементов. Из формулы (131) видно, что точка $z = 0$ является регулярной особой точкой для системы (132). Ряд (128) формально удовлетворяет этой системе. Поэтому, применяя лемму, заключаем, что ряд (128) сходится в окрестности $|z| < r$ точки $z = 0$.

В частности, сходится и ряд для $A(z)$ в формуле (110).

Таким образом, нами доказана следующая

Теорема 2. *Всякая система*

$$\frac{dX}{dz} = P(z)X \tag{133}$$

с регулярной особой точкой $z = 0$:

$$P(z) = \frac{P_{-1}}{z} + \sum_{m=0}^{\infty} P_m z^m$$

имеет решение вида

$$X = A(z)\, z^M z^U, \tag{134}$$

где $A(z)$ — *матричная функция, регулярная при* $z=0$ *и обращающаяся в этой точке в единичную матрицу* E, *а* M *и* U — *постоянные матрицы, причем* M *имеет простую структуру и целые характеристические числа, и разность между любыми двумя различными характеристическими числами матрицы* U *не есть целое число.*

Если матрица P_{-1} *приводится к жордановой форме при помощи неособенной матрицы* T

$$P_{-1} = T\{\lambda_1 E_1 + H_1,\ \lambda_2 E_2 + H_2,\ \ldots,\ \lambda_s E_s + H_s\}\, T^{-1} \tag{135}$$
$$(\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_s),$$

то можно взять M *и* U *в виде*

$$M = T\{m_1 E_1,\ m_2 E_2,\ \ldots,\ m_s E_s\}\, T^{-1}, \tag{136}$$

$$U = T\begin{pmatrix} \tilde\lambda_1 E_1 + H_1 & B_{12} & \ldots & B_{1s} \\ 0 & \tilde\lambda_2 E_2 + H_2 & \ldots & B_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & \tilde\lambda_s E_s + H_s \end{pmatrix} T^{-1}, \tag{137}$$

где

$$m_i = [\lambda_i], \quad \tilde\lambda_i = \lambda_i - m_i \quad (i = 1, 2, \ldots, s), \tag{138}$$

B_{ik} — *правильные нижние треугольные матрицы* $(i,\ k = 1,\ 2,\ \ldots,\ s)$, *причем* $B_{ik} = 0$, *если* $\lambda_i - \lambda_k$ *не есть целое положительное число* $(i, k = 1, 2, \ldots, s)$.

В частном случае, когда ни одна из разностей $\lambda_i - \lambda_k$ $(i, k = 1, 2, \ldots, s)$ *не равна целому положительному числу, в формуле* (134) *можно положить* $M = 0$, $U = P_{-1}$, *т. е. в этом случае решение представимо в виде*

$$X = A(z)\, z^{P_{-1}}. \tag{139}$$

Замечание 1. Обращаем внимание на то, что в настоящем параграфе был установлен алгоритм для определения коэффициентов ряда $A(z) = \sum_{m=0}^{\infty} A_m z^m$ $(A_0 = E)$ через коэффициенты P_m ряда для $P(z)$. Кроме того, доказанная теорема определяет и интегральную подстановку V, на которую умножается решение (134) при однократном обходе особой точки $z = 0$ в положительном направлении:

$$V = e^{2\pi i U}.$$

Замечание 2. Из формулировки теоремы следует, что

$$B_{ik} = 0 \ \text{ при } \ \tilde\lambda_i \neq \tilde\lambda_k \quad (i,\ k = 1, 2, \ldots, s).$$

Поэтому матрицы

$$\tilde\Lambda = T\{\tilde\lambda_1 E_1,\ \tilde\lambda_2 E_2,\ \ldots,\ \tilde\lambda_s E_s\}\, T^{-1} \ \text{ и } \ \tilde U = T\begin{pmatrix} 0 & B_{12} & \ldots & B_{1s} \\ 0 & 0 & \ldots & B_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 0 \end{pmatrix} T^{-1} \tag{140}$$

перестановочны между собой:

$$\widetilde{\Lambda}\widetilde{U}=\widetilde{U}\widetilde{\Lambda}.$$

Отсюда

$$z^{M}z^{U}=z^{M}z^{\widetilde{\Lambda}+\widetilde{U}}=z^{M}z^{\widetilde{\Lambda}}z^{\widetilde{U}}=z^{\Lambda}z^{\widetilde{U}}, \tag{141}$$

где

$$\Lambda=M+\widetilde{\Lambda}=T\{\lambda_1, \lambda_2, \ldots, \lambda_n\}T^{-1}, \tag{142}$$

где $\lambda_1, \lambda_2, \ldots, \lambda_n$ — все характеристические числа матрицы P_{-1}, расположенные в порядке $\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_n$.

С другой стороны,

$$z^{\widetilde{U}}=h(\widetilde{U}),$$

где $h(\lambda)$ — интерполяционный многочлен Лагранжа-Сильвестра для функции $f(\lambda)=z^{\lambda}$.

Поскольку все характеристические числа матрицы $\widetilde{U}$ равны нулю, то $h(\lambda)$ линейно зависит от $f(0), f'(0), \ldots, f^{(g-1)}(0)$, т. е. от $1, \ln z, \ldots, (\ln z)^{g-1}$ (g — наименьший показатель, при котором $\widetilde{U}^{g}=0$). Поэтому

$$h(\lambda)=\sum_{j=0}^{g-1}h_j(\lambda)(\ln z)^j$$

и потому

$$z^{\widetilde{U}}=h(\widetilde{U})=\sum_{j=0}^{g-1}h_j(\widetilde{U})(\ln z)^j=T\begin{pmatrix}1 & q_{12} & \ldots & q_{1n}\\ 0 & 1 & \ldots & q_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & 1\end{pmatrix}T^{-1}, \tag{143}$$

где q_{ij} ($i, j=1, 2, \ldots, n$; $i<j$) — многочлены от $\ln z$ степени ниже g.

В силу (134), (141), (142) и (143) частное решение системы (126) можно взять в виде

$$X=A(z)\begin{Vmatrix}z^{\lambda_1} & 0 & \ldots & 0\\ 0 & z^{\lambda_2} & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & z^{\lambda_n}\end{Vmatrix}\begin{Vmatrix}1 & q_{12} & \ldots & q_{1n}\\ 0 & 1 & \ldots & q_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & 1\end{Vmatrix}. \tag{144}$$

Здесь $\lambda_1, \lambda_2, \ldots, \lambda_n$ — характеристические числа матрицы P_{-1}, расположенные в порядке $\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_n$, а q_{ij} ($i, j=1, 2, \ldots, n$; $i<j$) — многочлены от $\ln z$ степени не выше $g-1$, где g — максимальное количество характеристических чисел λ_i, отличающихся между собой на целое число; $A(z)$ — матричная функция, регулярная в точке $z=0$, причем $A(0)=T$ ($|T|\neq 0$). Если матрица P_{-1} имеет жорданову форму, то $T=E$.

§ 11. Приводимые аналитические системы

В качестве приложения теоремы предыдущего параграфа выясним, в каких случаях система

$$\frac{dX}{dt}=Q(t)X, \tag{145}$$

где

$$Q(t)=\sum_{m=1}^{\infty}\frac{Q_m}{t^m} \tag{146}$$

— сходящийся ряд при $t>t_0$, является приводимой (по Ляпунову), т. е. в каких случаях существует решение системы вида

$$X=L(t)\,e^{Bt}, \tag{147}$$

где $L(t)$ — матрица Ляпунова (т. е. $L(t)$ удовлетворяет условиям 1° — 3° на стр. 422), а B — постоянная матрица[1]). Здесь X, Q — матрицы с комплексными элементами, а t — вещественный аргумент.

Сделаем преобразование

$$z=\frac{1}{t}.$$

Тогда система (145) перепишется в виде

$$\frac{dX}{dz}=P(z)\,X, \tag{148}$$

где

$$P(z)=-z^{-2}Q\left(\frac{1}{z}\right)=-\frac{Q_1}{z}-\sum_{m=0}^{\infty}Q_{m+2}z^m. \tag{149}$$

Ряд, стоящий в правой части выражения для $P(z)$, сходится при $|z|<\frac{1}{t_0}$. Могут представиться два случая:

1) $Q_1=0$. В этом случае точка $z=0$ не является особой для системы (148). Эта система имеет решение, регулярное и нормированное в точке $z=0$. Это решение задается сходящимся степенным рядом

$$X(z)=E+X_1z+X_2z^2+\ldots \qquad \left(|z|<\frac{1}{t_0}\right).$$

Полагая

$$L(t)=X\left(\frac{1}{t}\right), \quad B=0,$$

получим искомое представление (147). Система приводима.

2) $Q_1\neq 0$. В этом случае система (148) имеет регулярную особую точку в точке $z=0$.

Не нарушая общности рассуждений, можно считать матрицу вычетов $P_{-1}=-Q_1$ приведенной к жордановой форме, в которой диагональные элементы $\lambda_1, \lambda_2, \ldots, \lambda_n$ расположены в порядке $\operatorname{Re}\lambda_1\geqslant\operatorname{Re}\lambda_2\geqslant\ldots\geqslant\operatorname{Re}\lambda_n$.

Тогда в формуле (144) $T=E$, и потому система (148) имеет решение

$$X=A(z)\begin{Vmatrix} z^{\lambda_1} & 0 & \ldots & 0\\ 0 & z^{\lambda_2} & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & z^{\lambda_n}\end{Vmatrix}\begin{Vmatrix} 1 & q_{12} & \ldots & q_{1n}\\ 0 & 1 & \ldots & q_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & 1\end{Vmatrix},$$

---

[1]) Если имеет место равенство (147), то преобразование Ляпунова $X=L(t)\,Y$ переводит систему (145) в систему $\frac{dY}{dt}=BY$.

где функция $A(z)$ регулярна при $z=0$ и принимает в этой точке значение E, а q_{ik} $(i, k=1, 2, \ldots, n;\ i<k)$ — многочлены от $\ln z$. Заменяя здесь z на $\frac{1}{t}$, будем иметь:

$$X=A\left(\frac{1}{t}\right)\begin{Vmatrix}\left(\frac{1}{t}\right)^{\lambda_1} & 0 & \ldots & 0\\ 0 & \left(\frac{1}{t}\right)^{\lambda_2} & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & \left(\frac{1}{t}\right)^{\lambda_n}\end{Vmatrix}\begin{Vmatrix}1 & q_{12}\left(\ln\frac{1}{t}\right) & \ldots & q_{1n}\left(\ln\frac{1}{t}\right)\\ 0 & 1 & \ldots & q_{2n}\left(\ln\frac{1}{t}\right)\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & 1\end{Vmatrix}. \tag{150}$$

Так как преобразование $X=A\left(\frac{1}{t}\right)Y$ является преобразованием Ляпунова, то система (145) будет приводимой к некоторой системе с постоянными коэффициентами в том и только в том случае, когда произведение

$$L_1(t)=\begin{Vmatrix}t^{-\lambda_1} & 0 & \ldots & 0\\ 0 & t^{-\lambda_2} & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & t^{-\lambda_n}\end{Vmatrix}\begin{Vmatrix}1 & q_{12}\left(\ln\frac{1}{t}\right) & \ldots & q_{1n}\left(\ln\frac{1}{t}\right)\\ 0 & 1 & \ldots & q_{2n}\left(\ln\frac{1}{t}\right)\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & 1\end{Vmatrix}e^{-Bt}, \tag{151}$$

где B — некоторая постоянная матрица будет матрицей Ляпунова, т. е. когда матрицы $L_1(t)$, $\frac{dL_1}{dt}$ и $L_1^{-1}(t)$ будут ограничены. При этом, как следует из теоремы Еругина (§ 4), матрицу B можно считать матрицей с вещественными характеристическими числами.

Из ограниченности матриц $L_1(t)$ и $L_1^{-1}(t)$ при $t>t_0$ вытекает, что все характеристические числа матрицы B должны равняться нулю. Это следует из выражения для e^{Bt} и e^{-Bt}, получаемого из (151). Кроме того, все числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ должны быть чисто мнимыми, поскольку согласно (151) из ограниченности элементов последней строки в $L_1(t)$ и первого столбца в $L_1^{-1}(t)$ вытекает, что $\operatorname{Re}\lambda_n\geqslant 0$ и $\operatorname{Re}\lambda_1\leqslant 0$.

Но если все характеристические числа матрицы P_{-1} чисто мнимы, то разность между любыми двумя различными характеристическими числами матрицы P_{-1} не равна целому числу. Поэтому имеет место формула (139):

$$X=A(z)\,z^{P_{-1}}=A\left(\frac{1}{t}\right)t^{Q_1},$$

и для приводимости системы необходимо и достаточно, чтобы матрица

$$L_2(t)=t^{Q_1}e^{-Bt} \tag{152}$$

вместе со своей обратной была бы ограничена при $t>t_0$.

Поскольку все характеристические числа матрицы B должны равняться нулю, то минимальный многочлен для матрицы B имеет вид λ^d. Обозначим через

$$\psi(\lambda)=(\lambda-\mu_1)^{c_1}(\lambda-\mu_2)^{c_2}\ \ldots\ (\lambda-\mu_u)^{c_u}\qquad (\mu_i\neq\mu_k\ \text{ при }\ i\neq k)$$

минимальный многочлен матрицы Q_1. Поскольку $Q_1 = -P_{-1}$, то числа μ_1, μ_2, ..., μ_u отличаются знаком от соответствующих чисел λ_i и потому все они — чисто мнимые числа. Тогда [см. формулы (12), (13) на стр. 421]

$$t^{Q_1} = \sum_{k=1}^{u} [U_{k0} + U_{k1} \ln t + \ldots + U_{k,\, c_k - 1} (\ln t)^{c_k - 1}]\, t^{\mu_k}, \tag{153}$$

$$e^{Bt} = V_0 + V_1 t + \ldots + V_{d-1} t^{d-1}. \tag{154}$$

Подставляя эти выражения в равенство

$$L_2(t)\, e^{Bt} = t^{Q_1},$$

получим:

$$[L_2(t)\, V_{d-1} + (*)]\, t^{d-1} = Z_0(t) (\ln t)^{c-1}, \tag{155}$$

где c — наибольшее из чисел c_1, c_2, ..., c_n, $(*)$ обозначает матрицу, стремящуюся к нулю при $t \to \infty$, а $Z_0(t)$ — ограниченная матрица при $t > t_0$.

Так как матрицы, стоящие в левой и правой частях равенства (155), должны иметь одинаковый порядок роста при $t \to \infty$, то

$$d = c = 1,$$

т. е.

$$B = 0,$$

и матрица Q_1 имеет простые элементарные делители.

Обратно, если матрица Q_1 имеет простые элементарные делители и чисто мнимые характеристические числа μ_1, μ_2, ..., μ_n, то

$$X = A(z)\, z^{-Q_1} = A(z) \,\|\, z^{-\mu_i} \delta_{ik} \,\|_1^n$$

есть решение системы (149). Полагая здесь $z = \frac{1}{t}$, найдем:

$$X = A\left(\frac{1}{t}\right) \| \, t^{\mu_i} \delta_{ik} \, \|_1^n.$$

Функция $X(t)$ вместе с $\frac{dX(t)}{dt}$ и обратной матрицей $X^{-1}(t)$ ограничена при $t > t_0$. Поэтому система приводима ($B = 0$). Нами доказана[1])

Теорема 3. *Система*

$$\frac{dX}{dt} = Q(t)\, X,$$

где матрица $Q(t)$ представима сходящимся при $t > t_0$ рядом

$$Q(t) = \frac{Q_1}{t} + \frac{Q_2}{t^2} + \ldots,$$

является приводимой в том и только в том случае, если у матрицы вычетов Q_1 все элементарные делители простые и все характеристические числа чисто мнимы.

---

[1]) См. работу Еругина [11], стр. 21—23. Там эта теорема доказана для случая, когда матрица Q_1 не имеет различных характеристических чисел, отличающихся между собой на целое число.

§ 12. Аналитические функции от многих матриц и их применение к исследованию дифференциальных систем. Работы И. А. Лаппо-Данилевского

Аналитическая функция от m матриц n-го порядка $X_1, X_2, \ldots, X_m$ может быть задана при помощи ряда

$$F(X_1, X_2, \ldots, X_m) = \alpha_0 + \sum_{\nu=1}^{\infty} \sum_{j_1, j_2, \ldots, j_\nu}^{(1 \ldots m)} \alpha_{j_1 j_2 \ldots j_\nu} X_{j_1} X_{j_2} \ldots X_{j_\nu}, \tag{156}$$

сходящегося для всех матриц m-го порядка X_j, удовлетворяющих неравенствам

$$\operatorname{mod} X_j < R_j \qquad (j = 1, 2, \ldots, m). \tag{157}$$

Здесь коэффициенты

$$\alpha_0, \alpha_{j_1 j_2 \ldots j_\nu} \qquad (j_1, j_2, \ldots, j_\nu = 1, 2, \ldots, m; \ \nu = 1, 2, 3, \ldots)$$

— комплексные числа, $R_j (j = 1, 2, \ldots, m)$ — постоянные матрицы n-го порядка с положительными элементами и $X_j (j = 1, 2, \ldots, m)$ — матрицы того же порядка, но переменные и с комплексными элементами.

Теория аналитических функций от нескольких матриц была развита И. А. Лаппо-Данилевским. На основе этой теории И. А. Лаппо-Данилевский провел фундаментальные исследования систем линейных дифференциальных уравнений с рациональными коэффициентами.

Система с рациональными коэффициентами путем надлежащего преобразования независимой переменной всегда может быть приведена к виду

$$\frac{dV}{dz} = \sum_{j=1}^{m} \left\{ \frac{U_{j0}}{(z-a_j)^{s_j}} + \frac{U_{j1}}{(z-a_j)^{s_j-1}} + \ldots + \frac{U_{j, s_j-1}}{z-a_j} \right\} X, \tag{158}$$

где U_{jk} — постоянные матрицы n-го порядка, a_j — комплексные числа, s_j — целые положительные числа ($k = 0, 1, \ldots, s_j - 1$; $j = 1, 2, \ldots, m$)[1].

Некоторые результаты Лаппо-Данилевского мы проиллюстрируем на частном случае так называемых *регулярных* систем. Последние характеризуются условием $s_1 = s_2 = \ldots = s_m = 1$ и записываются в виде

$$\frac{dX}{dz} = \sum_{j=1}^{m} \frac{U_j}{z-a_j} X. \tag{159}$$

Следуя Лаппо-Данилевскому, введем в рассмотрение специальные аналитические функции — гиперлогарифмы, — определяемые следующими рекуррентными соотношениями:

$$l_b(z; a_{j_1}) = \int_b^z \frac{dz}{z-a_{j_1}},$$

$$l_b(z; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) = \int_b^z \frac{l_b(z; a_{j_2}, a_{j_3}, \ldots, a_{j_\nu})}{z-a_{j_1}} dz.$$

[1]) В системе (158) все коэффициенты — правильные рациональные дроби относительно z. К такому виду приводятся любые рациональные коэффициенты, если при помощи дробно-линейного преобразования над переменной z перевести регулярную (для всех коэффициентов) конечную точку $z = c$ в точку $z = \infty$.

Рассматривая точки $a_1, a_2, \ldots, a_m, \infty$ как точки разветвления логарифмического типа, построим соответствующую риманову поверхность $S(a_1, a_2, \ldots, a_m; \infty)$. Каждый гиперлогарифм будет однозначной функцией на этой поверхности. С другой стороны, матрицант системы (159) Ω_b^z (т. е. нормированное в точке $z=b$ решение), будучи аналитически продолжен, также может быть рассматриваем как однозначная функция на $S(a_1, a_2, \ldots, a_m; \infty)$; при этом в качестве b может быть выбрана любая конечная точка на S, отличная от $a_1, a_2, \ldots, a_m$.

Для нормированного решения Ω_b^z Лаппо-Данилевский дает явное выражение через определяющие матрицы $U_1, U_2, \ldots, U_m$ системы (159) в виде ряда

$$\Omega_b^z = E + \sum_{\nu=1}^{\infty} \sum_{j_1, \ldots, j_\nu}^{(1\ldots m)} l_b(z; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) U_{j_1} U_{j_2} \ldots U_{j_\nu}. \tag{160}$$

Это разложение сходится равномерно относительно z при любых $U_1, U_2, \ldots, U_m$ и представляет Ω_b^z в любой конечной области на поверхности $S(a_1, a_2, \ldots, a_m; \infty)$, если только эта область не содержит внутри и на границе точек $a_1, \ldots, a_m$.

Если ряд (156) сходится при любых матрицах $X_1, X_2, \ldots, X_m$, то соответствующая функция $F(X_1, X_2, \ldots, X_m)$ называется *целой*. Ω_b^z представляет собой целую функцию от матриц $U_1, U_2, \ldots, U_m$.

Заставляя в формуле (160) аргумент z обойти точку a_j в положительном направлении один раз так, чтобы контур обхода не захватывал других точек a_i (при $i \neq j$), мы получим выражение для *интегральной подстановки* V_j, соответствующей точке $z = a_j$:

$$V_j = E + \sum_{\nu=1}^{\infty} \sum_{j_1, \ldots, j_\nu}^{(1\ldots m)} p_j(b; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) U_{j_1} U_{j_2} \ldots U_{j_\nu} \tag{161}$$

$$(j = 1, 2, \ldots, m),$$

где в понятных обозначениях

$$p_j(b; a_{j_1}) = \int_{(a_j)} \frac{\partial z}{z - a_{j_1}},$$

$$p_j(b; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) = \int_{(a_j)} \frac{L_b(z; a_{j_2}, a_{j_3}, \ldots, a_{j_\nu})}{z - a_{j_1}} dz$$

$$\begin{pmatrix} j_1, j_2, \ldots, j_\nu, \; j = 1, 2, \ldots, m; \\ \nu = 1, 2, 3, \ldots \end{pmatrix}.$$

Ряд (161), как и ряд (160), представляет собой целую функцию от $U_1, U_2, \ldots, U_m$.

Обобщив теорию аналитических функций на случай бесконечного, но счетного множества матриц-аргументов $X_1, X_2, X_3, \ldots$,[1]). Лаппо-Данилевский использовал эту теорию для исследования поведения решения системы в окрестности иррегулярной особой точки[2]). Мы приведем основной результат.

---

[1]) См. [21б], т. I, Мемуар 1.

[2]) См. [21б], т. I, Мемуар 3, см. также [92], [104а, б].

Нормированное решение Ω_b^z системы

$$\frac{dX}{dz}=\sum_{j=-q}^{+\infty} P_j z^j X,$$

где степенной ряд в правой части сходится при $|z|<r\,(r>1)$[1]), может быть представлено рядом

$$\Omega_b^z=E+\sum_{\nu=1}^{\infty}\sum_{j_1,\,j_2,\ldots,j_\nu=-q}^{\infty} P_{j_1}\ldots$$

$$\ldots P_{j_\nu}\sum_{\mu=0}^{\nu} b^{j_{\mu+1}+\ldots+j_\nu+\nu-\mu} z^{j_1+\ldots+j_\mu+\mu}\sum_{\lambda=0}^{n-\mu} \alpha^{*(\lambda)}_{j_{\mu+1},\ldots,j_\nu}\ln^\lambda b\sum_{\varkappa=0}^{\mu}\alpha^{(\varkappa)}_{j_1,\ldots,j_\mu}\ln^\varkappa z. \tag{162}$$

Здесь $\alpha^{*(\lambda)}_{j_{\mu+1},\ldots,j_\nu}$ и $\alpha^{(\varkappa)}_{j_1,\ldots,j_\mu}$ — скалярные коэффициенты, определяемые по специальным формулам. Ряд (162) сходится при любых матрицах P_1, $P_2,\ldots$ в кольце

$$\varrho<|z|<r$$

(ϱ — произвольное положительное число, меньшее r). Этому кольцу должна принадлежать и точка $b\,(\varrho<|b|<r)$.

Не имея возможности в какой бы то ни было степени подробно изложить содержание работ Лаппо-Данилевского в настоящей книге, мы вынуждены ограничиться приведенными выше формулировками некоторых основных результатов и отослать читателей к соответствующей литературе.

Все относящиеся к дифференциальным уравнениям работы Лаппо-Данилевского изданы посмертно Академией наук СССР в трех томах в 1934—1936 гг. Кроме того, основные результаты автора изложены в статьях [82] и небольшой книге [18а]. Сокращенное изложение некоторых результатов можно найти и в книге В. И. Смирнова «Курс высшей математики», т. III.

---

[1]) Ограничение $r>1$ несущественно, так как это условие всегда можно получить заменой z на αz, где α — надлежащим образом выбранное положительное число.

ГЛАВА XVI

ПРОБЛЕМА РАУСА — ГУРВИЦА И СМЕЖНЫЕ ВОПРОСЫ

§ 1. Введение

В главе XII, § 2 мы выяснили, что согласно теореме Ляпунова нулевое решение системы дифференциальных уравнений

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} a_{ik}x_k + (**) \qquad (i = 1, \ldots, n) \tag{1}$$

[a_{ik} (i, $k = 1, 2, \ldots, n$) — постоянные коэффициенты] при любых членах ($**$) второго порядка и выше относительно $x_1, x_2, \ldots, x_n$ является устойчивым, если все характеристические числа матрицы $A = \| a_{ik} \|_1^n$, т. е. все корни векового уравнения $\Delta(\lambda) \equiv | \lambda E - A | = 0$, имеют отрицательные вещественные части.

Поэтому задача установления необходимых и достаточных условий, при которых все корни данного алгебраического уравнения расположены в левой полуплоскости, имеет фундаментальное значение в ряде прикладных областей, в которых исследуется устойчивость механических и электрических систем.

Важность этой алгебраической задачи была ясна основоположникам теории регулирования машин, английскому физику Д. К. Максвеллу и русскому инженеру-исследователю И. А. Вышнеградскому, которые в своих работах, посвященных регуляторам, установили и широко использовали упомянутые алгебраические условия для уравнений не выше третьей степени [1]).

В 1868 г. Максвелл выдвинул математическую задачу об отыскании соответствующих условий для алгебраического уравнения любой степени. Между тем эта задача по существу была решена в опубликованной в 1856 г. работе французского математика Эрмита [128]. В этой работе была установлена тесная связь между числом корней комплексного многочлена $f(z)$, расположенных внутри какой-либо полуплоскости (или даже внутри какого-либо прямоугольника), и сигнатурой некоторой квадратичной формы. Однако результаты Эрмита не были доведены до такого состояния, чтобы они могли быть использованы специалистами, работающими в прикладных областях. Поэтому эта работа Эрмита и не получила соответствующего распространения.

---

[1]) Д. К. Максвелл, О регуляторах (1868); И. А. Вышнеградский, О регуляторах прямого действия (1876). Эти работы опубликованы в сборнике «Теория автоматического регулирования» (издание АН СССР, 1949). См. там же статью А. А. Андронова и И. Н. Вознесенского, О работах Д. К. Максвелла, И. А. Вышнеградского и А. Стодолы в области теории регулирования машин.

В 1875 г. английский механик Раус [45], пользуясь теоремой Штурма и теорией индексов Коши, установил алгоритм для определения числа k корней вещественного многочлена, расположенных в правой полуплоскости ($\operatorname{Re} z > 0$). В частном случае $k = 0$ этот алгоритм и дает критерий устойчивости.

В конце XIX в. крупнейший словацкий инженер-исследователь, создатель теории паровых и газовых турбин, А. Стодола, не зная работы Рауса, снова поставил задачу об отыскании условий того, чтобы все корни алгебраического уравнения имели отрицательные вещественные части, и в 1895 г. А. Гурвиц [129], опираясь на работы Эрмита, дает второе (независимое от Рауса) решение той же задачи. Полученные Гурвицем детерминантные неравенства известны в настоящее время под названием условий Рауса — Гурвица.

Однако еще до появления в свет работы Гурвица основатель современной теории устойчивости А. М. Ляпунов в своей знаменитой диссертации («Общая задача об устойчивости движения», Харьков, 1892) установил [1]) теорему, из которой вытекают необходимые и достаточные условия для того, чтобы все корни характеристического уравнения вещественной матрицы $A = \| a_{ik} \|_1^n$ имели отрицательные вещественные части. Эти условия используются в ряде работ по теории регулирования [2]).

Новый критерий устойчивости был установлен в 1914 г. французскими математиками Льенаром и Шипаром [135].

Используя специальные квадратичные формы, эти авторы получили критерий устойчивости, имеющий некоторые преимущества перед критерием Рауса — Гурвица (число детерминантных неравенств в критерии Льенара — Шипара примерно вдвое меньше, нежели в критерии Рауса — Гурвица).

Знаменитые русские математики П. Л. Чебышев и А. А. Марков установили две замечательные теоремы в связи с разложением в ряды непрерывных дробей специального типа. Эти теоремы, как будет показано в § 16, имеют непосредственное отношение к проблеме Рауса — Гурвица.

В очерченном круге вопросов, как увидит читатель, находят себе существенное применение теория квадратичных форм (гл. X) и, в частности, теория ганкелевых форм (гл. X, § 10).

§ 2. Индексы Коши

Начнем с рассмотрения так называемых индексов Коши [3]).

Определение 1. *Индексом Коши* вещественной рациональной функции $R(x)$ в пределах от a до b (обозначение $I_a^b R(x)$; a, b — вещественные числа, либо $\pm\infty$) будем называть разность между числом разрывов $R(x)$ с переходом от $-\infty$ к $+\infty$ и числом разрывов с переходом от $+\infty$ к $-\infty$ при изменении аргумента от a к b [4]).

Согласно этому определению, если

$$R(x) = \sum_{i=1}^{p} \frac{A_i}{x - \alpha_i} + R_1(x),$$

[1]) См. [22], § 20.

[2]) См., например, [73].

[3]) См. [10], стр. 419—425.

[4]) При подсчете числа разрывов крайние значения x — пределы a и b — не включаются.

где A_i, α_i $(i = 1, 2, \ldots, p)$ — вещественные числа, а $R_1(x)$ — рациональная функция, не имеющая вещественных полюсов [1]), то

$$I_{-\infty}^{+\infty} R(x) = \sum_{i=1}^{p} \operatorname{sign} A_i \text{ [2]} \tag{2}$$

и вообще

$$I_a^b R(x) = \sum_{a<\alpha_i<b} \operatorname{sign} A_i \quad (a < b). \tag{2'}$$

В частности, если $f(x) = a_0(x-\alpha_1)^{n_1} \ldots (x-\alpha_m)^{n_m}$ — вещественный многочлен ($\alpha_i \neq \alpha_k$ при $i \neq k$; $i, k = 1, 2, \ldots, m$) и среди корней $\alpha_1, \alpha_2, \ldots, \alpha_m$ этого многочлена только первые p вещественны, то

$$\frac{f'(x)}{f(x)} = \sum_{j=1}^{m} \frac{n_j}{x-\alpha_j} = \sum_{i=1}^{p} \frac{n_i}{x-\alpha_i} + R_1(x),$$

где $R_1(x)$ — вещественная рациональная функция, не имеющая вещественных полюсов.

Поэтому *индекс*

$$I_a^b \frac{f'(x)}{f(x)} \quad (a < b)$$

равен числу различных вещественных корней многочлена $f(x)$, *находящихся внутри интервала* (a, b).

Произвольная вещественная рациональная функция $R(x)$ всегда представима в виде

$$R(x) = \sum_{i=1}^{p} \left\{ \frac{A_1^{(i)}}{x-\alpha_i} + \ldots + \frac{A_{n_i}^{(i)}}{(x-\alpha_i)^{n_i}} \right\} + R_1(x),$$

где все α и A — вещественные числа ($A_{n_i}^{(i)} \neq 0$; $i = 1, 2, \ldots, p$) и $R_1(x)$ не имеет вещественных полюсов.

Тогда

$$I_{-\infty}^{+\infty} R(x) = \sum_{(n_i \text{ нечетно})} \operatorname{sign} A_{n_i}^{(i)} \tag{3}$$

и вообще

$$I_a^b R(x) = \sum_{\substack{a<\alpha_i<b \\ n_i \text{ нечетно}}} \operatorname{sign} A_{n_i}^{(i)} \quad (a < b) \text{ [3]}. \tag{3'}$$

Если $R(a) = R(b) = 0$, то индекс $I_a^b R(x)$ выражается через приращение непрерывной функции $\operatorname{arctg} R(x)$:

$$I_a^b R(x) = -\Delta_a^b \operatorname{arctg} R(x) \quad (a < b) \text{ [4]}. \tag{4}$$

---

[1]) Полюсами рациональной функции являются те значения аргумента, при которых эта функция обращается в бесконечность.

[2]) Под $\operatorname{sign} a$ (a — вещественное число) мы понимаем $+1$, -1 или 0 в зависимости от того, $a > 0$, $a < 0$ или $a = 0$.

[3]) В (3) сумма распространяется на все те значения i, для которых соответствующее n_i нечетно. В (3′) сумма распространяется на все те i, для которых n_i нечетно и $a < \alpha_i < b$.

[4]) Если $a = -\infty$, а $b = +\infty$, то формула (4) справедлива для любой правильной рациональной дроби $R(x)$, поскольку в этом случае $R(-\infty) = R(+\infty) = 0$.

Один из методов вычисления индекса $I_a^b R(x)$ основан на классической теореме Штурма.

Рассмотрим ряд вещественных многочленов

$$f_1(x),\ f_2(x),\ \ldots,\ f_m(x), \tag{5}$$

обладающий двумя свойствами по отношению к интервалу $(a,\ b)$ [1]:

1° При любом значении $x\ (a < x < b)$, обращающем в нуль какую-либо из функций $f_k(x)$, две смежные функции $f_{k-1}(x)$ и $f_{k+1}(x)$ имеют значения, отличные от нуля и разных знаков, т. е. из $f_k(x) = 0$ при $a < x < b$ следует: $f_{k-1}(x) f_{k+1}(x) < 0$.

2° Последняя функция $f_m(x)$ в ряду (5) не обращается в нуль внутри $(a,\ b)$, т. е. $f_m(x) \neq 0$ при $a < x < b$.

Такой ряд (4) многочленов называется *рядом Штурма в интервале* $(a,\ b)$.

Обозначим через $V(x)$ число перемен знака в ряду (5) при фиксированном значении x [2]. Тогда при изменении x от a до b величина $V(x)$ может измениться лишь при переходе через нуль какой-либо из функций ряда (5). Но в силу 1° при переходе через нуль функции $f_k(x)$ $(k = 2, \ldots, m-1)$ величина $V(x)$ не изменяется. При переходе же через нуль функции $f_1(x)$ теряется или приобретается одна перемена знака в ряду (5) в зависимости от того, переходит при этом отношение $\frac{f_2(x)}{f_1(x)}$ от $-\infty$ к $+\infty$ или наоборот. Поэтому имеет место

Теорема 1 (Штурма). *Если $f_1(x),\ f_2(x),\ \ldots,\ f_m(x)$ — ряд Штурма в $(a,\ b)$, а $V(x)$ — число перемен знака в этом ряду, то*

$$I_a^b \frac{f_2(x)}{f_1(x)} = V(a) - V(b). \tag{6}$$

Примечание. Помножим все члены ряда Штурма на один и тот же произвольный многочлен $d(x)$. Полученный таким образом ряд многочленов назовем *обобщенным рядом Штурма*. Так как умножение всех членов ряда (5) на один и тот же многочлен не меняет ни левой, ни правой части равенства (6), то теорема Штурма сохраняет свою силу и для обобщенного ряда Штурма.

Заметим, что если даны два произвольных многочлена $f(x)$ и $g(x)$ [степень $f(x) \geqslant$ степени $g(x)$], то всегда можно при помощи алгоритма Евклида построить обобщенный ряд Штурма, который начинался бы с функций $f_1(x) \equiv f(x)$, $f_2(x) \equiv g(x)$.

Действительно, обозначая через $-f_3(x)$ остаток от деления $f_1(x)$ на $f_2(x)$, через $-f_4(x)$ — остаток от деления $f_2(x)$ на $f_3(x)$ и т. д., будем иметь цепочку тождеств

$$\begin{aligned} f_1(x) &= q_1(x) f_2(x) - f_3(x),\ \ldots,\ f_{k-1}(x) = \\ &= q_{k-1}(x) f_k(x) - f_{k+1}(x),\ \ldots,\ f_{m-1}(x) = q_{m-1}(x) f_m(x), \end{aligned} \tag{7}$$

где последний не равный тождественно нулю остаток $f_m(x)$ является наибольшим общим делителем $f(x)$ и $g(x)$, а также наибольшим общим

---

[1]) При этом a может равняться $-\infty$, а b $+\infty$.

[2]) Если $a < x < b$ и $f_1(x) \neq 0$, то в силу 1° при определении $V(x)$ нулевые значения в ряду (4) можно опустить либо этим значениям можно приписать произвольные знаки. Если a конечно, то под $V(a)$ следует понимать $V(a + \varepsilon)$, где ε — столь малое положительное число, что в полузамкнутом интервале $(a,\ a + \varepsilon]$ ни одна из функций $f_i(x)$ $(i = 1, 2, \ldots, m)$ не обращается в нуль. Точно так же, если b конечно, то под $V(b)$ следует понимать $V(b - \varepsilon)$, где число ε определяется аналогично.

делителем всех функций построенного таким образом ряда (5). Если $f_m(x) \neq 0$ $(a < x < b)$, то полученный ряд (5) в силу (7) удовлетворяет условиям 1°, 2° и является рядом Штурма. Если же многочлен $f_m(x)$ имеет корни внутри интервала (a, b), то ряд (5) является обобщенным рядом Штурма, поскольку он становится рядом Штурма после деления всех его членов на $f_m(x)$.

Из сказанного следует, что индекс любой рациональной функции $R(x)$ может быть определен при помощи теоремы Штурма. Для этого достаточно представить $R(x)$ в виде $Q(x) + \frac{g(x)}{f(x)}$, где $Q(x)$, $f(x)$, $g(x)$ — многочлены и степень $g(x) \leqslant$ степени $f(x)$. Тогда, если построить обобщенный ряд Штурма для $f(x)$, $g(x)$, то

$$I_a^b R(x) = I_a^b \frac{g(x)}{f(x)} = V(a) - V(b).$$

При помощи теоремы Штурма можно определить число различных вещественных корней многочлена $f(x)$ внутри интервала (a, b), поскольку это число, как мы видели, равно $I_a^b \frac{f'(x)}{f(x)}$.

§ 3. Алгоритм Рауса

1. Задача Рауса состоит в определении числа k корней вещественного многочлена $f(z)$, расположенных в правой полуплоскости $(\operatorname{Re} z > 0)$.

Рассмотрим сначала случай, когда $f(z)$ не имеет нулей на мнимой оси. В правой полуплоскости построим полуокружность радиуса R с центром в нуле и рассмотрим область, ограниченную этой полуокружностью и отрезком мнимой оси (рис. 9). При достаточно большом R все k нулей многочлена $f(z)$ с положительными вещественными частями будут находиться внутри этой области. Поэтому $\arg f(z)$ при положительном обходе контура области получит приращение $2k\pi$[1]). С другой стороны, приращение $\arg f(z)$ вдоль полуокружности радиуса R при $R \to \infty$ определяется приращением аргумента старшего члена $a_0 z^n$ и потому равно $n\pi$. Поэтому для приращения $\arg f(z)$ вдоль мнимой оси $(R \to \infty)$ получаем выражение

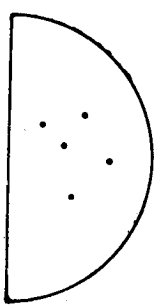

Рис. 9.

$$\Delta_{-\infty}^{+\infty} \arg f(i\omega) = (n - 2k)\pi. \tag{8}$$

Введем не совсем обычные обозначения для коэффициентов многочлена $f(z)$, пусть

$$f(z) = a_0 z^n + b_0 z^{n-1} + a_1 z^{n-2} + b_1 z^{n-3} + \ldots \quad (a_0 \neq 0).$$

Тогда, замечая, что приращение $\Delta \arg f(i\omega)$ в формуле (8) не изменится, если многочлен $f(z)$ помножить на произвольное комплексное число, положим:

$$\frac{1}{i^n} f(i\omega) = f_1(\omega) - i f_2(\omega), \tag{9}$$

---

[1]) В самом деле, если $f(z) = a_0 \prod_{i=1}^{n} (z - z_i)$, то $\Delta \arg f(z) = \sum_{i=1}^{n} \Delta \arg (z - z_i)$. Если точка z_i находится внутри рассматриваемой области, то $\Delta \arg (z - z_i) = 2\pi$; если z_i вне этой области, то $\Delta \arg (z - z_i) = 0$.

где

$$\left.\begin{aligned} f_1(\omega) &= a_0\omega^n - a_1\omega^{n-2} + a_3\omega^{n-4} - \ldots, \\ f_2(\omega) &= b_0\omega^{n-1} - b_1\omega^{n-3} + b_3\omega^{n-5} - \ldots \end{aligned}\right\} \tag{10}$$

Следуя Раусу, воспользуемся индексом Коши. Из формул (4) и (9) находим:

$$\frac{1}{\pi}\Delta_{-\infty}^{+\infty}\arg f(i\omega) = -\frac{1}{\pi}\Delta_{-\infty}^{+\infty}\operatorname{arctg}\frac{f_2(\omega)}{f_1(\omega)} = I_{-\infty}^{+\infty}\frac{f_2(\omega)}{f_1(\omega)}.$$

Поэтому из формулы (8) следует, что[1])

$$I_{-\infty}^{+\infty}\frac{b_0\omega^{n-1} - b_1\omega^{n-3} + \ldots}{a_0\omega^n - a_1\omega^{n-2} + \ldots} = n - 2k. \tag{11}$$

2. Для определения индекса, стоящего в левой части равенства (11), используем теорему Штурма (см. предыдущий параграф). Исходя из функций $f_1(\omega)$ и $f_2(\omega)$, определяемых равенствами (10), построим, следуя Раусу, при помощи алгоритма Евклида обобщенный ряд Штурма (см. стр. 471)

$$f_1(\omega),\ f_2(\omega),\ f_3(\omega),\ \ldots,\ f_m(\omega). \tag{12}$$

Рассмотрим теперь регулярный случай: $m = n+1$. В этом случае в ряду (12) степень каждой функции на единицу меньше степени предыдущей, и последняя функция $f_m(\omega)$ имеет нулевую степень[2]).

Из алгоритма Евклида [см. (7)] следует, что

$$f_3(\omega) = \frac{a_0}{b_0}\omega f_2(\omega) - f_1(\omega) = c_0\omega^{n-2} - c_1\omega^{n-4} + c_2\omega^{n-6} - \ldots,$$

где

$$c_0 = a_1 - \frac{a_0}{b_0}b_1 = \frac{b_0a_1 - a_0b_1}{b_0},\quad c_1 = a_2 - \frac{a_0}{b_0}b_2 = \frac{b_0a_2 - a_0b_2}{b_0},\ \ldots \tag{13}$$

Точно так же

$$f_4(\omega) = \frac{b_0}{c_0}\omega f_3(\omega) - f_2(\omega) = d_0\omega^{n-3} - d_1\omega^{n-5} + \ldots,$$

где

$$d_0 = b_1 - \frac{b_0}{c_0}c_1 = \frac{c_0b_1 - b_0c_1}{c_0},\qquad d_1 = b_2 - \frac{b_0}{c_0}c_2 = \frac{c_0b_2 - b_0c_2}{c_0},\ \ldots \tag{13'}$$

Аналогично определяются коэффициенты остальных многочленов $f_5(\omega), \ldots, f_{n+1}(\omega)$.

При этом каждый из многочленов

$$f_1(\omega),\ f_2(\omega),\ \ldots,\ f_{n+1}(\omega) \tag{14}$$

является четной или нечетной функцией, причем степени смежных многочленов всегда имеют разную четность.

---

[1]) Напомним, что формула (10) выведена в предположении, что многочлен $f(z)$ не имеет корней на мнимой оси.

[2]) В регулярном случае ряд (12) является обычным (необобщенным) рядом Штурма.

Составим схему Рауса:

$$\left.\begin{array}{llll} a_0, & a_1, & a_2, & \ldots, \\ b_0, & b_1, & b_2, & \ldots, \\ c_0, & c_1, & c_2, & \ldots, \\ d_0, & d_1, & d_2, & \ldots, \\ \cdot & \cdot & \cdot & \cdot \end{array}\right\} \tag{15}$$

В этой схеме, как показывают формулы (13), (13′), каждая строка определяется из двух предыдущих по следующему правилу:

Из чисел верхней строки вычитаются соответствующие числа нижней, предварительно помноженные на такое число, чтобы первая разность обращалась в нуль. Отбрасывая эту нулевую разность, получаем искомую строку.

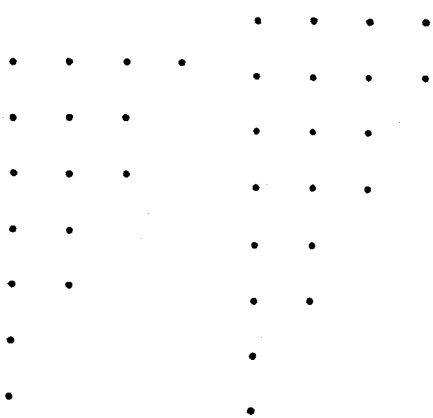

Рис. 10. Рис. 11.

Регулярный случай, очевидно, характеризуется тем, что при последовательном применении этого правила мы в ряду

$$b_0,\ c_0,\ d_0,\ \ldots$$

не встречаем числа, равного нулю, и этот ряд состоит из $n+1$ чисел.

На рис. 10 и 11 показан скелет регулярной схемы Рауса при n четном ($n=6$) и n нечетном ($n=7$). Здесь элементы схемы отмечены точками.

В регулярном случае многочлены $f_1(\omega)$ и $f_2(\omega)$ имеют наибольший общий делитель $f_{n+1}(\omega)=\text{const}\neq 0$. Поэтому эти многочлены не обращаются одновременно в нуль, т. е. $f(i\omega)=f_1(\omega)-if_2(\omega)\neq 0$ при ω вещественном. Поэтому *в регулярном случае имеет место формула* (11).

Применяя к левой части этой формулы теорему Штурма в интервале $(-\infty,\ +\infty)$ и используя при этом ряд (14), получаем согласно (11):

$$V(-\infty)-V(+\infty)=n-2k. \tag{16}$$

В данном случае[1])

$$V(+\infty)=V(a_0,\ b_0,\ c_0,\ d_0,\ \ldots),$$

а

$$V(-\infty)=V(a_0,\ -b_0,\ c_0,\ -d_0,\ \ldots).$$

Отсюда

$$V(-\infty)+V(+\infty)=n. \tag{17}$$

Из равенств (16) и (17) находим:

$$k=V(a_0,\ b_0,\ c_0,\ d_0,\ \ldots). \tag{18}$$

Нами доказана для регулярного случая

[1]) Знак $f_k(\omega)$ при $\omega=+\infty$ совпадает со знаком старшего коэффициента, а при $\omega=-\infty$ отличается от этого знака множителем $(-1)^{n-k+1}$ $(k=1,\ 2,\ \ldots,\ n+1)$.

Теорема 2 (Рауса). *Число корней вещественного многочлена $f(z)$, лежащих в правой полуплоскости* $\operatorname{Re} z > 0$, *равно числу перемен знака в первом столбце схемы Рауса.*

3. Рассмотрим важный частный случай, когда все корни $f(z)$ имеют отрицательные вещественные части («случай устойчивости»). В этом случае многочлен $f(z)$ не имеет чисто мнимых корней, и потому имеет место формула (11), а следовательно, и формула (16). Поскольку $k=0$, формула (16) перепишется так:

$$V(-\infty)-V(+\infty)=n. \tag{19}$$

Но $0 \leqslant V(-\infty) \leqslant m-1 \leqslant n$ и $0 \leqslant V(+\infty) \leqslant m-1 \leqslant n$. Поэтому равенство (19) возможно лишь тогда, когда $m=n+1$ (регулярный случай!) и $V(+\infty)=0$, $V(-\infty)=m-1=n$. Тогда из формулы (18) следует:

Критерий Рауса. *Для того чтобы все корни вещественного многочлена $f(z)$ имели отрицательные вещественные части, необходимо и достаточно, чтобы при выполнении алгоритма Рауса все элементы первого столбца схемы Рауса получались отличными от нуля и одного знака.*

4. При установлении теоремы Рауса мы опирались на формулу (11). В дальнейшем нам понадобится обобщение этой формулы. Формула (11) была выведена в предположении, что многочлен $f(z)$ не имеет корней на мнимой оси. Мы покажем, что в общем случае, когда многочлен $f(z)=$ $=a_0z^n+b_0z^{n-1}+a_1z^{n-2}+\ldots$ $(a_0 \neq 0)$ имеет k корней в правой полуплоскости и s корней на мнимой оси, формула (11) заменяется формулой

$$I_{-\infty}^{+\infty} \frac{b_0\omega^{n-1}-b_1\omega^{n-3}+b_2\omega^{n-5}-\ldots}{a_0\omega^n-a_1\omega^{n-2}+a_2\omega^{n-4}-\ldots}=n-2k-s. \tag{20}$$

В самом деле,

$$f(z)=d(z)\,f^*(z),$$

где вещественный многочлен $d(z)=z^s+\ldots$ имеет s корней на мнимой оси, а многочлен $f^*(z)$ степени $n^*=n-s$ таких корней не имеет. Пусть

$$\frac{1}{i^n} f(i\omega)=f_1(\omega)-if_2(\omega), \qquad \frac{1}{i^{n-s}} f^*(i\omega)=f_1^*(i\omega)-if_2^*(i\omega).$$

Тогда

$$f_1(\omega)-if_2(\omega)=\frac{1}{i^s}\,d(i\omega)\,[f_1^*(i\omega)-if_2^*(i\omega)].$$

Поскольку $\frac{1}{i^s}\,d(i\omega)$ — вещественный многочлен относительно ω, то

$$\frac{f_2(\omega)}{f_1(\omega)}=\frac{f_2^*(\omega)}{f_1^*(\omega)}.$$

К многочлену $f^*(z)$ применима формула (11). Поэтому

$$I_{-\infty}^{+\infty} \frac{f_2(\omega)}{f_1(\omega)}=I_{-\infty}^{+\infty} \frac{f_2^*(\omega)}{f_1^*(\omega)}=n^*-2k=n-2k-s,$$

что и требовалось доказать[1]).

---

[1]) Обращаем внимание читателя на интересное обобщение критерия Рауса, содержащееся в работе Фаэдо [176]. Здесь устанавливаются достаточные условия того, что корни всех многочленов $f(z)=a_0z^n+a_1z^{n-1}+\ldots+a_n$ с коэффициентами a_i, меняющимися в заданных интервалах

$$\underline{a}_i \leqslant a_i \leqslant \bar{a}_i \qquad (i=0,\ 1,\ \ldots,\ n;\quad \underline{a}_0<0),$$

одновременно все имеют отрицательные вещественные части.

§ 4. Особые случаи. Примеры

1. В предыдущем параграфе мы разобрали регулярный случай, когда при заполнении схемы Рауса ни одно из чисел b_0, c_0, d_0, ... не оказывается равным нулю.

Переходим теперь к рассмотрению особых случаев, когда в ряду чисел b_0, c_0, ... мы встречаемся с числом $h_0=0$. Алгоритм Рауса останавливается на той строке, где находится h_0, так как для получения чисел следующей строки нужно делить на h_0.

Особые случаи могут быть двух типов:

1) В той же строке, где находится h_0, имеются числа, не равные нулю. Это означает, что в каком-то месте ряда (12) произошло понижение степени больше чем на единицу.

2) Одновременно все числа строки, содержащей h_0, оказываются равными нулю. Тогда эта строка является $(m+1)$-й, где m — число членов в обобщенном ряду Штурма (5). В этом случае в ряду (12) степени функций все время понижаются на единицу, но степень последней функции $f_m(\omega)$ больше нуля. В обоих случаях в ряду (12) число функций $m<n+1$.

Поскольку обычный алгоритм Рауса в особых случаях приостанавливается, Раус дает специальные правила для продолжения схемы в случаях 1), 2).

2. В случае 1) следует по Раусу вместо $h_0=0$ подставить «малую» величину ε определенного (но произвольного) знака и продолжать заполнение схемы. При этом последующие элементы первого столбца схемы будут рациональными функциями от ε. Знаки этих элементов определятся, исходя из «малости» и знака ε. Если же какой-либо из этих элементов окажется тождественным нулем относительно ε, то мы этот элемент заменим другой малой величиной η и продолжим алгоритм.

Пример.

$$f(z)=z^4+z^3+2z^2+2z+1.$$

Схема Рауса (с малым параметром ε)

$$\begin{array}{ll} 1, & 2, \quad 1 \\ 1, & 2 \\ \varepsilon. & 1 \\ 2-\dfrac{1}{\varepsilon} & \\ 1 & \end{array} \qquad k=V\left(1,\ 1,\ \varepsilon,\ 2-\frac{1}{\varepsilon},\ 1\right)=2.$$

Обоснование этого своеобразного метода варьирования элементов схемы заключается в следующем:

Поскольку мы предполагаем отсутствие особенностей второго типа, то функции $f_1(\omega)$ и $f_2(\omega)$ взаимно просты. Отсюда следует, что многочлен $f(z)$ не имеет корней на мнимой оси.

В схеме Рауса все элементы выражаются рационально через элементы первых двух строк, т. е. через коэффициенты данного многочлена. Но нетрудно усмотреть из формул (13), (13′) и аналогичных формул для последующих строк, что, задавшись произвольными значениями для элементов двух любых подряд идущих строк схемы Рауса и для первых элементов предыдущих строк, мы можем целым рациональным образом выразить через эти элементы все числа, стоящие в первых двух строках, т. е. коэффициенты исходного многочлена. Так, например, все числа a, b можно представить в виде целых рациональных функций от

$$a_0,\ b_0,\ c_0,\ \ldots,\ h_0,\ h_1,\ h_2,\ \ldots,\ g_0,\ g_1,\ g_2,\ \ldots$$

Поэтому, заменяя $h_0=0$ на ε, мы фактически видоизменяем наш исходный многочлен. Вместо схемы для $f(z)$ мы имеем схему Рауса для многочлена $F(z, \varepsilon)$, где $F(z, \varepsilon)$ — целая рациональная функция от z и ε, обращающаяся в $f(z)$ при $\varepsilon=0$. Так как корни многочлена $F(z, \varepsilon)$ непрерывно меняются с изменением параметра ε и при $\varepsilon=0$ нет корней на мнимой оси, то при малых по модулю значениях ε число k корней в правой полуплоскости у многочленов $F(z, \varepsilon)$ и $F(z, 0)=f(z)$ одинаково.

3. Переходим к рассмотрению особенностей второго типа. Пусть в схеме Рауса

$$a_0\neq 0,\ b_0\neq 0,\ \ldots,\ e_0\neq 0,\ h_0=0,\ h_1=0,\ h_2=0,\ \ldots$$

В этом случае в обобщенном ряду Штурма (16) последний многочлен имеет вид:

$$f_m(\omega)=e_0\omega^{n-m+1}-e_1\omega^{n-m-1}+\ldots$$

Раус предлагает заменить нулевое $f_{m+1}(\omega)$ на $f'_m(\omega)$, т. е. вместо нулевых h_0, h_1, ... написать соответственно коэффициенты

$$(n-m+1)\,e_0,\quad (n-m-1)\,e_1,\ \ldots$$

и продолжать алгоритм.

Обоснование этого правила заключается в следующем:

Согласно формуле (20)

$$I_{-\infty}^{+\infty}\frac{f_2(\omega)}{f_1(\omega)}=n-2k-s$$

s корней многочлена $f(z)$ на мнимой оси совпадают с вещественными корнями многочлена $f_m(\omega)$. Поэтому, если эти вещественные корни простые, то (см. стр. 470)

$$I_{-\infty}^{+\infty}\frac{f'_m(\omega)}{f_m(\omega)}=s,$$

и, следовательно,

$$I_{-\infty}^{+\infty}\frac{f_2(\omega)}{f_1(\omega)}+I_{-\infty}^{+\infty}\frac{f'_m(\omega)}{f_m(\omega)}=n-2k.$$

Эта формула показывает, что недостающую часть схемы Рауса следует заполнить схемой Рауса для многочленов $f_m(\omega)$ и $f'_m(\omega)$. Коэффициенты многочлена $f'_m(\omega)$ и используются для замены элементов нулевой строки в схеме Рауса.

Если же корни $f_m(\omega)$ не простые, то, обозначая через $d(\omega)$ наибольший общий делитель $f_m(\omega)$ и $f'_m(\omega)$, через $e(\omega)$ наибольший общий делитель $d(\omega)$ и $d'(\omega)$ и т. д., мы будем иметь:

$$I_{-\infty}^{+\infty}\frac{f'(\omega)}{f(\omega)}+I_{-\infty}^{+\infty}\frac{d'(\omega)}{d(\omega)}+I_{-\infty}^{+\infty}\frac{e'(\omega)}{e(\omega)}+\ldots=s.$$

Таким образом, искомое число k можно получить, если недостающую часть схемы Рауса дополнить схемами Рауса для $f_m(\omega)$ и $f'_m(\omega)$, $d(\omega)$ и $d'(\omega)$, $e(\omega)$ и $e'(\omega)$ и т. д., т. е. несколько раз применять правило Рауса для ликвидации особенностей 2-го типа.

Пример.

$$f(z)=z^{10}+z^9-z^8-2z^7+z^6+3z^5+z^4-2z^3-z^2+z+1.$$

Схема

ω^{10}	1	−1	1	1	−1	1
ω^9	1	−2	3	−2	1	
ω^8	1	−2	3	−2	1	
ω^7	8	−12	12	−4		
	2	−3	3	−1		
ω^6	−1	3	−3	2		
ω^5	3	−3	3			
	1	−1	1			
ω^4	2	−2	2			
	1	−1	1			
ω^3	4	−2				
	2	−1				
ω^2	−1	2				
ω	1					
ω^0	2					
	1					

$$k=V(1,\ 1,\ 1,\ 2,\ -1,\ 1,\ 1,\ 2,\ -1,\ 1,\ 1)=4.$$

Примечание. Не изменяя знаков элементов первого столбца, можно все элементы какой-либо строки помножить на одно и то же число. Это замечание было использовано при построении схемы.

4. Однако применение обоих правил Рауса не дает возможности во всех случаях определить число k. Применение первого правила (введение малых параметров ε, η, ...) обосновано лишь в том случае, когда многочлен $f(z)$ не имеет корней на мнимой оси.

Если многочлен $f(z)$ имеет корни на мнимой оси, то при варьировании параметра ε некоторые из этих корней могут перейти в правую полуплоскость и изменить число k.

Пример.

$$f(z)=z^6+z^5+3z^4+3z^3+3z^2+2z+1.$$

Схема

ω^6	1	3	3	1
ω^5	1	3	2	
ω^4	ε	1	1	
ω^3	$3-\frac{1}{\varepsilon}$	$2-\frac{1}{\varepsilon}$		$\left(2-\frac{1}{\varepsilon}-\frac{3-\frac{1}{\varepsilon}}{1-\frac{2\varepsilon-1}{3-\frac{1}{\varepsilon}}}=-\varepsilon+\ldots\right)$
ω^2	$1-\frac{2\varepsilon-1}{3-\frac{1}{\varepsilon}}$	1		
ω	u			$V\left(1,\ 1,\ \varepsilon,\ 3-\frac{1}{\varepsilon},\ 1,\ -\varepsilon,\ 1\right)=\begin{cases}4 \text{ при } \varepsilon>0,\\ 2 \text{ при } \varepsilon<0.\end{cases}$
ω^0	1			

Вопрос, чему равно k, остается открытым.

В общем случае, когда $f(z)$ имеет корни на мнимой оси, следует поступать так: Полагая $f(z)=F_1(z)+F_2(z)$, где

$$F_1(z)=a_0z^n+a_1z^{n-2}+\ldots,\quad F_2(z)=b_0z^{n-1}+b_1z^{n-3}+\ldots,$$

следует найти наибольший общий делитель $d(z)$ многочленов $F_1(z)$ и $F_2(z)$. Тогда $f(z)=d(z)f^*(z)$.

Если $f(z)$ имеет корень z, для которого $-z$ снова является корнем (этим свойством обладают и все корни на мнимой оси), то из $f(z)=0$ и $f(-z)=0$ следует: $F_1(z)=0$ и $F_2(z)=0$, т. е. z является корнем $d(z)$. Поэтому многочлен $f^*(z)$ не имеет корней z, для которых $-z$ является корнем $f^*(z)$ [1]).

Тогда

$$k=k_1+k_2,$$

где k_1 и k_2 — числа корней в правой полуплоскости многочленов $f^*(z)$ и $d(z)$; k_1 определяется по алгоритму Рауса, а $k_2=\frac{q-s}{2}$, где q — степень $d(z)$, а s — число вещественных корней многочлена $d(i\omega)$ [2]).

В последнем примере

$$d(z)=z^2+1,\quad f^*(z)=z^4+z^3+2z^2+2z+1.$$

Поэтому (см. пример на стр. 476) здесь $k_2=0$, $k_1=2$ и, следовательно,

$$k=2.$$

---

[1]) При определении многочлена $d(z)$ можно исходить не из функций $F_1(z)$ и $F_2(z)$, а из введенных ранее (стр. 472) функций $f_1(z)$ и $f_2(z)$. (Подробнее об этом см. сноску к стр. 490.)

[2]) $d(i\omega)$ — вещественный многочлен или становится таковым после сокращения на i. Число вещественных корней его можно определить при помощи теоремы Штурма.

§ 5. Теорема Ляпунова

Из исследований А. М. Ляпунова, опубликованных в 1892 г. в его монографии «Общая задача об устойчивости движения», вытекает теорема [1]), дающая необходимые и достаточные условия для того, чтобы все корни характеристического уравнения $|\lambda E - A| = 0$ вещественной матрицы $A = \| a_{ik} \|_1^n$ имели отрицательные вещественные части. Поскольку любой многочлен $f(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_n \ (a_0 \neq 0)$ может быть представлен в виде характеристического определителя $|\lambda E - A|$ [2]), то теорема Ляпунова носит общий характер и относится к любому алгебраическому уравнению $f(\lambda) = 0$.

Пусть дана вещественная матрица $A = \| a_{ik} \|_1^n$ и однородный многочлен m-го измерения относительно переменных $x_1, x_2, \ldots, x_n$

$$V(\underbrace{x, x, \ldots, x}_{m}) \qquad [x = (x_1, x_2, \ldots, x_n)].$$

Найдем полную производную по t от функции $V(x, x, \ldots, x)$ в предположении, что x есть решение дифференциальной системы $\frac{dx}{dt} = Ax$. Тогда

$$\frac{d}{dt} V(x, x, \ldots, x) = V(Ax, x, \ldots, x) +$$
$$+ V(x, Ax, \ldots, x) + \ldots + V(x, x, \ldots, Ax) = W(x, x, \ldots, x), \qquad (21)$$

где $W(x, x, \ldots, x)$ — снова однородный многочлен m-го измерения относительно $x_1, x_2, \ldots, x_n$. Равенство (21) определяет линейный оператор $\widehat{A}$, относящий каждому однородному многочлену m-го измерения $V(x, x, \ldots, x)$ некоторый однородный многочлен $W(x, x, \ldots, x)$ того же измерения m

$$W = \widehat{A}(V).$$

Мы ограничимся случаем $m = 2$ [3]). В этом случае $V(x, x)$ и $W(x, x)$ — квадратичные формы от переменных $x_1, x_2, \ldots, x_n$, связанные равенством

$$\frac{d}{dt} V(x, x) = V(Ax, x) + V(x, Ax) = W(x, x), \qquad (22)$$

---

[1]) См. [19], § 20.

[2]) Для этого достаточно, например, положить:

$$A = \begin{Vmatrix} 0 & 0 & \ldots & 0 & -\frac{a_n}{a_0} \\ 1 & 0 & \ldots & 0 & -\frac{a_{n-1}}{a_0} \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & \ddots & \\ & & & \ddots & \\ 0 & 0 & \ldots & 1 & -\frac{a_1}{a_0} \end{Vmatrix}.$$

[3]) А. М. Ляпунов установил свою теорему (см. ниже теорему 3) при любом целом и положительном m.

откуда[1])

$$W = \widehat{A}(V) = A'V + VA. \tag{23}$$

Здесь $V = \| v_{ik} \|_1^n$ и $W = \| w_{ik} \|_1^n$ — симметрические матрицы, составленные соответственно из коэффициентов форм $V(x, x)$ и $W(x, x)$. Линейный оператор $\widehat{A}$ в пространстве симметрических матриц n-го порядка V целиком определяется заданием матрицы $A = \| a_{ik} \|_1^n$.

Если $\lambda_1, \lambda_2, \ldots, \lambda_n$ — характеристические числа матрицы A, то каждое характеристическое число оператора $\widehat{A}$ представляется в виде $\lambda_i + \lambda_k$ $(1 \leqslant i \leqslant k \leqslant n)$.

Действительно, пусть u_k — собственный вектор-столбец матрицы A', соответствующий характеристическому числу λ_k, т. е. $Au_k = \lambda_k u_k$ $(u_k \neq 0)$, и пусть $v_{ik} = u_i u_k'$. Тогда

$$\widehat{A} v_{ik} = A' u_i u_k' + u_i u_k' A = (A' u_i) u_k' + u_i (A' u_k)' =$$

$$= (\lambda_i + \lambda_k) u_i u_k' = (\lambda_i + \lambda_k) v_{ik} \qquad (i, k = 1, \ldots, n). \tag{23'}$$

Если все величины $(\lambda_i + \lambda_k)$ $(i, k = 1, \ldots, n)$ различны между собой, то из равенств (23′) следует, что эти величины образуют полную систему характеристических чисел оператора $\widehat{A}$. Общий случай, когда среди сумм $\lambda_i + \lambda_k$ имеются равные между собой, получается из рассмотренного случая с помощью соображений непрерывности.

Из доказанного предложения следует, что оператор $\widehat{A}$ является неособенным, матрица $A = \| a_{ik} \|_1^n$ не имеет нулевых и двух противоположных характеристических чисел. В этом случае задание матрицы W однозначно определяет матрицы V в (23).

Таким образом, *если матрица $A = \| a_{ik} \|_1^n$ не имеет нулевых и двух противоположных характеристических чисел, то каждой квадратичной форме $W(x, x)$ отвечает одна и только одна квадратичная форма $V(x, x)$, связанная с $V(x, x)$ равенством* (22).

Теперь сформулируем теорему Ляпунова.

Теорема 3 (Ляпунова). *Если все характеристические числа вещественной матрицы $A = \| a_{ik} \|_1^n$ имеют отрицательные вещественные части, то любой отрицательно определенной квадратичной форме $W(x, x)$ отвечает положительно определенная квадратичная форма $V(x, x)$, связанная с формой $W(x, x)$ в силу уравнения*

$$\frac{dx}{dt} = Ax \tag{24}$$

равенством

$$\frac{d}{dt} V(x, x) = W(x, x). \tag{25}$$

Обратно, если для некоторой отрицательно определенной формы $W(x, x)$ существует положительно определенная форма $V(x, x)$, связанная с $W(x, x)$ равенством (25) *в силу уравнения* (24), *то все характеристические числа матрицы $A = \| a_{ik} \|_1^n$ имеют отрицательные вещественные части.*

---

[1]) Поскольку $V(x, y) = x'Vy$.

Доказательство. 1. Пусть все характеристические числа матрицы A имеют отрицательные вещественные части. Тогда для любого решения $x=e^{At}x_0$ системы (24) имеем: $\lim\limits_{t\to+\infty} x=0$[1]). Пусть формы $V(x, x)$ и $W(x, x)$ связаны формулой (25) и $W(x, x)<0$ $(x\neq 0)$[2]).

Допустим, что при некотором $x_0\neq 0$

$$V_0=V(x_0, x_0)\leqslant 0.$$

Но $\frac{d}{dt}V(x, x)=W(x, x)<0\,(x=e^{At}x_0)$. Поэтому при $t>0$ величина $V(x, x)$ отрицательна и убывает при $t\to+\infty$, что находится в противоречии с равенством $\lim\limits_{t\to+\infty} V(x, x)=\lim\limits_{x\to 0} V(x, x)=0$. Следовательно, $V(x, x)>0$ при $x\neq 0$, т. е. $V(x, x)$ — положительно определенная квадратичная форма.

2. Пусть, обратно, дано, что в равенстве (25)

$$W(x, x)<0,\ V(x, x)>0 \qquad (x\neq 0).$$

Из (25) следует:

$$V(x, x)=V(x_0, x_0)+\int\limits_0^t W(x, x)\,dt \qquad (x=e^{At}x_0). \tag{25'}$$

Докажем, что при произвольном $x_0\neq 0$ столбец $x=e^{At}x_0$ как угодно близко подходит к нулю при некоторых сколь угодно больших значениях $t>0$. Допустим противное. Тогда существует число $\nu>0$ такое, что

$$W(x, x)<-\nu<0 \qquad (x=e^{At}x_0,\ x_0\neq 0,\ t>0).$$

Но тогда из (25′)

$$V(x, x)<V(x_0, x_0)-\nu t,$$

и, следовательно, при некоторых достаточно больших значениях t справедливо неравенство $V(x, x)<0$, что противоречит условию.

Из доказанного следует, что при некоторых достаточно больших значениях t величина $V(x, x)\,(x=e^{At}x_0,\ x_0\neq 0)$ будет как угодно близка к нулю. Но $V(x, x)$ монотонно убывает при $t>0$, поскольку $\frac{d}{dt}V(x, x)=$ $=W(x, x)<0$. Поэтому $\lim\limits_{t\to\infty} V(x, x)=0$.

Отсюда вытекает, что при любом $x_0\neq 0$ имеет место равенство $\lim\limits_{t\to+\infty} e^{At}x_0=0$, т. е. $\lim\limits_{t\to+\infty} e^{At}=0$. Это возможно лишь тогда, когда все характеристические числа матрицы A имеют отрицательные вещественные части (см. гл. V, § 6).

Теорема доказана полностью.

В качестве формы $W(x, x)$ в теореме Ляпунова можно взять любую отрицательно определенную форму и, в частности, форму $-\sum\limits_{i=1}^{n} x_i^2$. В этом случае теорема допускает следующую матричную формулировку:

---

[1]) См. гл. V, § 6.

[2]) Форма $W(x, x)$ нам произвольно задана. Форма $V(x, x)$ однозначно определяется из условия (25), поскольку в данном случае матрица A не имеет нулевых и двух противоположных характеристических чисел.

Теорема 3'. *Для того чтобы все характеристические числа вещественной матрицы* $A = \| a_{ik} \|_1^n$ *имели отрицательные вещественные части, необходимо и достаточно, чтобы матричное уравнение*

$$A'V + VA = -E \tag{26}$$

имело в качестве решения V *матрицу коэффициентов некоторой положительно определенной квадратичной формы* $V(x, x) > 0$.

Из доказанной теоремы вытекает критерий для определения устойчивости нелинейной системы по ее линейному приближению[1]).

Пусть требуется доказать асимптотическую устойчивость нулевого решения нелинейной системы дифференциальных уравнений (1) (стр. 419) в том случае, когда коэффициенты a_{ik} $(i, k = 1, 2, \ldots, n)$ в линейных членах правых частей уравнений образуют матрицу $A = \| A_{ik} \|_1^n$, имеющую только характеристические числа с отрицательными вещественными частями. Тогда, определяя положительно определенную форму $V(x, x)$ при помощи матричного уравнения (26) и вычисляя ее полную производную по времени в предположении, что $x = (x_1, x_2, \ldots, x_n)$ есть решение системы (1), будем иметь:

$$\frac{d}{dt} V(x, x) = -\sum_{i=1}^{n} x_i^2 + R(x_1, x_2, \ldots, x_n),$$

где $R(x_1, x_2, \ldots, x_n)$ — ряд, содержащий члены третьего и более высоких измерений относительно $x_1, x_2, \ldots, x_n$. Поэтому в некоторой достаточно малой окрестности точки $(0, 0, \ldots, 0)$ для любого $x \neq 0$ одновременно

$$V(x, x) > 0, \qquad \frac{d}{dt} V(x, x) < 0.$$

Согласно общему критерию устойчивости Ляпунова[2]) это и означает асимптотическую устойчивость нулевого решения системы дифференциальных уравнений.

Если из матричного уравнения (26) выразить элементы матрицы V через элементы матрицы A и полученные выражения подставить в неравенства

$$v_{11} > 0, \qquad \begin{vmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{vmatrix} > 0, \quad \ldots, \quad \begin{vmatrix} v_{11} & v_{12} & \ldots & v_{1n} \\ v_{21} & v_{22} & \ldots & v_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ v_{n1} & v_{n2} & \ldots & v_{nn} \end{vmatrix} > 0,$$

то мы получим неравенства, которым должны удовлетворять элементы матрицы $A = \| a_{ik} \|_1^n$ для того, чтобы все характеристические числа матрицы имели отрицательные вещественные части. Однако в значительно более простом виде эти неравенства могут быть получены из критерия Рауса — Гурвица, которому посвящается следующий параграф.

---

[1]) См. [22], § 26; [38], стр. 113 и далее; [23], стр. 66 и далее.
[2]) См. [22], § 16; [38], стр. 19—21 и 31—33; [23], стр. 32—34.

Примечание. Теорема Ляпунова (3) или (3′) непосредственно обобщается на случай произвольной комплексной матрицы $A = \| a_{ik} \|_1^n$. В этом случае квадратичные формы $V(x, x)$ и $W(x, x)$ заменяются эрмитовыми

$$V(x, x) = \sum_{i, k=1}^{n} v_{ik} \bar{x}_i x_k, \qquad W(x, x) = \sum_{i, k=1}^{n} w_{ik} \bar{x}_i x_k.$$

В соответствии с этим матричное уравнение (26) заменится уравнением

$$A^* V + VA = -E \quad (A^* = \bar{A}').$$

§ 6. Теорема Рауса — Гурвица

В предыдущих параграфах был изложен непревзойденный по своей простоте метод Рауса для определения числа k корней в правой полуплоскости вещественного многочлена, коэффициенты которого заданы как конкретные числа. Если же коэффициенты многочлена зависят от параметров и требуется определить, при каких значениях параметров число k будет иметь то или другое значение и, в частности, значение 0 (область устойчивости!)[1]), то желательно иметь конкретные выражения для величин c_0, d_0, ... через коэффициенты данного многочлена. Разрешив эту задачу, мы получим метод определения числа k и, в частности, критерий устойчивости в том виде, в каком он был установлен Гурвицем [129].

Рассмотрим снова многочлен

$$f(z) = a_0 z^n + b_0 z^{n-1} + a_1 z^{n-2} + b_1 z^{n-3} + \ldots \quad (a_0 \neq 0).$$

Назовем *матрицей Гурвица* квадратную матрицу n-го порядка

$$H = \begin{Vmatrix} b_0 & b_1 & b_2 & \ldots & b_{n-1} \\ a_0 & a_1 & a_2 & \ldots & a_{n-1} \\ 0 & b_0 & b_1 & \ldots & b_{n-2} \\ 0 & a_0 & a_1 & \ldots & a_{n-2} \\ 0 & 0 & b_0 & \ldots & b_{n-3} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{Vmatrix} \quad \left(\begin{matrix} a_k = 0 \text{ при } k > \left[\frac{n}{2} \right], \\ b_k = 0 \text{ при } k > \left[\frac{n-1}{2} \right] \end{matrix} \right). \tag{27}$$

Преобразуем эту матрицу, вычитая из второй, четвертой, ... соответственно первую, третью, ... строки, предварительно помноженные на $\frac{a_0}{b_0}$[2]). Получим матрицу

$$\begin{Vmatrix} b_0 & b_1 & b_2 & \ldots & b_{n-1} \\ 0 & c_0 & c_1 & \ldots & c_{n-2} \\ 0 & b_0 & b_1 & \ldots & b_{n-2} \\ 0 & 0 & c_0 & \ldots & c_{n-3} \\ 0 & 0 & b_0 & \ldots & b_{n-3} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{Vmatrix}.$$

[1]) Так именно и обстоит дело при проектировании новых механических или электрических систем регулирования.

[2]) Сначала рассматривается регулярный случай, когда $b_0 \neq 0$, $c_0 \neq 0$, $d_0 \neq 0$, ...

Здесь $c_0, c_1, \ldots$ — третья строка схемы Рауса, дополненная нулями $\left(c_k = 0 \text{ при } k > \left[\frac{n}{2}\right] - 1\right)$.

Полученную матрицу снова преобразуем, вычитая из третьей, пятой, ... строки соответственно вторую, четвертую, ... строки, предварительно помноженные на $\frac{b_0}{c_0}$:

$$\begin{Vmatrix} b_0 & b_1 & b_2 & b_3 & \ldots \\ 0 & c_0 & c_1 & c_2 & \ldots \\ 0 & 0 & d_0 & d_1 & \ldots \\ 0 & 0 & c_0 & c_1 & \ldots \\ 0 & 0 & 0 & d_0 & \ldots \\ 0 & 0 & 0 & c_0 & \ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{Vmatrix}.$$

Продолжая этот процесс далее, мы придем в конце концов к треугольной матрице n-го порядка

$$R = \begin{Vmatrix} b_0 & b_1 & b_2 & \ldots \\ 0 & c_0 & c_1 & \ldots \\ 0 & 0 & d_0 & \ldots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{Vmatrix},$$

которую назовем *матрицей Рауса*. Она получается из схемы Рауса [см. (15)] 1) отбрасыванием первой строки, 2) сдвигом строк вправо так, чтобы их первые элементы пришлись на главную диагональ, и 3) пополнением нулями до квадратной матрицы n-го порядка.

Определение 2. Две матрицы $A = \| a_{ik} \|_1^n$ и $B = \| b_{ik} \|_1^n$ назовем *равносильными* в том и только в том случае, если при любом $p \leqslant n$ в первых p строках этих матриц соответствующие миноры p-го порядка равны между собой:

$$A\begin{pmatrix} 1 & 2 & \ldots & p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} = B\begin{pmatrix} 1 & 2 & \ldots & p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} \quad \begin{pmatrix} i_1, i_2, \ldots, i_p = 1, 2, \ldots, n, \\ p = 1, 2, \ldots, n \end{pmatrix}.$$

Так как при вычитании из какой-либо строки матрицы какой-либо предыдущей строки, помноженной предварительно на произвольное число, миноры p-го порядка в первых p строках $(p = 1, 2, \ldots, n)$ не меняют своей величины, то согласно определению 2 матрицы Гурвица и Рауса H и R равносильны:

$$H\begin{pmatrix} 1 & 2 & \ldots & p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} = R\begin{pmatrix} 1 & 2 & \ldots & p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} \quad \begin{pmatrix} i_1, i_2, \ldots, i_p = 1, 2, \ldots, n, \\ p = 1, 2, \ldots, n. \end{pmatrix}. \quad (28)$$

Равносильность матриц H и R позволяет выразить все элементы матрицы R, т. е. элементы схемы Рауса, через миноры матрицы Гурвица H и, следовательно, через коэффициенты данного многочлена. Действи-

тельно, давая p в (28) последовательно значения 1, 2, 3, ..., получим:

$$\left.\begin{array}{lll} H\begin{pmatrix}1\\1\end{pmatrix}=b_0, & H\begin{pmatrix}1\\2\end{pmatrix}=b_1, & H\begin{pmatrix}1\\3\end{pmatrix}=b_2, \ldots, \\ H\begin{pmatrix}1&2\\1&2\end{pmatrix}=b_0c_0, & H\begin{pmatrix}1&2\\1&3\end{pmatrix}=b_0c_1, & H\begin{pmatrix}1&2\\1&4\end{pmatrix}=b_0c_2, \ldots, \\ H\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}=b_0c_0d_0, & H\begin{pmatrix}1&2&3\\1&2&4\end{pmatrix}=b_0c_0d_1, & H\begin{pmatrix}1&2&3\\1&2&5\end{pmatrix}=b_0c_0d_2, \ldots \end{array}\right\} \quad (29)$$

и т. д.

Отсюда находим следующие выражения для элементов схемы Рауса:

$$\left.\begin{array}{lll} b_0=H\begin{pmatrix}1\\1\end{pmatrix}, & b_1=H\begin{pmatrix}1\\2\end{pmatrix}, & b_2=H\begin{pmatrix}1\\3\end{pmatrix}, \ldots, \\ c_0=\dfrac{H\begin{pmatrix}1&2\\1&2\end{pmatrix}}{H\begin{pmatrix}1\\1\end{pmatrix}}, & c_1=\dfrac{H\begin{pmatrix}1&2\\1&3\end{pmatrix}}{H\begin{pmatrix}1\\1\end{pmatrix}}, & c_2=\dfrac{H\begin{pmatrix}1&2\\1&4\end{pmatrix}}{H\begin{pmatrix}1\\1\end{pmatrix}}, \ldots, \\ d_0=\dfrac{H\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}}{H\begin{pmatrix}1&2\\1&2\end{pmatrix}}, & d_1=\dfrac{H\begin{pmatrix}1&2&3\\1&2&4\end{pmatrix}}{H\begin{pmatrix}1&2\\1&2\end{pmatrix}}, & d_2=\dfrac{H\begin{pmatrix}1&2&3\\1&2&5\end{pmatrix}}{H\begin{pmatrix}1&2\\1&2\end{pmatrix}}, \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\} \quad (30)$$

Последовательные главные миноры матрицы H обычно называются *определителями Гурвица*. Мы их будем обозначать через

$$\Delta_1=H\begin{pmatrix}1\\1\end{pmatrix}=b_0, \quad \Delta_2=H\begin{pmatrix}1&2\\1&2\end{pmatrix}=\begin{vmatrix}b_0&b_1\\a_0&a_1\end{vmatrix}, \ldots$$

$$\ldots, \Delta_n=H\begin{pmatrix}1&2\ldots n\\1&2\ldots n\end{pmatrix}=\begin{vmatrix}b_0&b_1\ldots b_{n-1}\\a_0&a_1\ldots a_{n-1}\\0&b_0\ldots b_{n-2}\\0&a_0\ldots a_{n-2}\\\ldots&\ldots\end{vmatrix}. \quad (31)$$

Замечание 1. Согласно формулам (29)

$$\Delta_1=b_0, \quad \Delta_2=b_0c_0, \quad \Delta_3=b_0c_0d_0, \ \ldots\ ^{1)}. \quad (32)$$

Из $\Delta_1\neq 0, \ldots, \Delta_p\neq 0$ следует, что первые p из числа $b_0, c_0, \ldots$ отличны от нуля и наоборот; в этом случае определены p подряд идущих строк схемы Рауса, начиная с третьей, и для них имеют место формулы (30).

Замечание 2. Регулярный случай (все $b_0, c_0, \ldots$ имеют смысл и не равны нулю) характеризуется неравенствами

$$\Delta_1\neq 0, \ \Delta_2\neq 0, \ \ldots, \ \Delta_n\neq 0.$$

Замечание 3. Определение элементов схемы Рауса при помощи формул (30) является более общим, нежели определение при помощи

---

[1]) Если коэффициенты многочлена $f(z)$ заданы численно, то формулы (32) дают наиболее простой способ вычисления определителей Гурвица, сводя это вычисление к составлению схемы Рауса.

алгоритма Рауса. Так, например, если $b_0 = H\begin{pmatrix}1\\1\end{pmatrix} = 0$, то алгоритм Рауса нам не дает ничего, кроме первых двух строк, составленных из коэффициентов данного многочлена. Однако, если при $\Delta_1 = 0$ остальные определители $\Delta_2, \Delta_3, \ldots$ отличны от нуля, мы при помощи формул (30), минуя строку из c, можем определить все последующие строки схемы Рауса.

Согласно формулам (32)

$$b_0 = \Delta_1, \quad c_0 = \frac{\Delta_2}{\Delta_1}, \quad d_0 = \frac{\Delta_3}{\Delta_2}, \quad \ldots,$$

и потому

$$V(a_0, b_0, c_0, \ldots) = V\left(a_0, \Delta_1, \frac{\Delta_2}{\Delta_1}, \ldots, \frac{\Delta_n}{\Delta_1}\right) =$$
$$= V(a_0, \Delta_1, \Delta_3, \ldots) + V(1, \Delta_2, \Delta_4, \ldots).$$

Поэтому теорема Рауса может быть сформулирована так:

Теорема 4 (Рауса—Гурвица). *Число k корней вещественного многочлена $f(z) = a_0 z^n + \ldots$, расположенных в правой полуплоскости, определяется формулой*

$$k = V\left(a_0, \Delta_1, \frac{\Delta_2}{\Delta_1}, \frac{\Delta_3}{\Delta_2}, \ldots, \frac{\Delta_n}{\Delta_{n-1}}\right) \tag{33}$$

или (что то же)

$$k = V(a_0, \Delta_1, \Delta_3, \ldots) + V(1, \Delta_2, \Delta_4, \ldots). \tag{33'}$$

Примечание. Приведенная формулировка теоремы Рауса—Гурвица предполагает, что имеет место регулярный случай

$$\Delta_1 \neq 0, \quad \Delta_2 \neq 0, \quad \ldots, \quad \Delta_n \neq 0.$$

В § 8 мы покажем, как пользоваться этой формулой в особых случаях, когда некоторые из определителей Гурвица Δ_i равны нулю.

Рассмотрим теперь тот частный случай, когда все корни многочлена $f(z)$ расположены в левой полуплоскости $\operatorname{Re} z < 0$. В этом случае согласно критерию Рауса все $a_0, b_0, c_0, d_0, \ldots$ должны быть отличны от нуля и одного знака. Так как здесь мы имеем дело с регулярным случаем, то получаем из (33) при $k = 0$ следующий критерий:

Теорема 5 (Критерий Рауса—Гурвица). *Для того чтобы у вещественного многочлена $f(z) = a_0 z^n + \ldots$ $(a_0 \neq 0)$ все корни имели отрицательные вещественные части, необходимо и достаточно, чтобы имели место неравенства*

$$a_0\Delta_1 > 0, \ \Delta_2 > 0, \ a_0\Delta_3 > 0, \ \Delta_4 > 0, \ \ldots, \ \left.\begin{matrix} a_0\Delta_n > 0 \text{ (при } n \text{ нечетном)}, \\ \Delta_n > 0 \text{ (при } n \text{ четном)}. \end{matrix}\right\} \tag{34}$$

Примечание. Если $a_0 > 0$, то эти условия записываются так:

$$\Delta_1 > 0, \quad \Delta_2 > 0, \quad \ldots, \quad \Delta_n > 0. \tag{35}$$

Если принять обычные обозначения для коэффициентов многочлена $f(z) = a_0 z^n - a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_{n-1} z + a_n$, то при $a_0 > 0$ условия Рауса—Гурвица (35) записываются в виде следующих детерминантных

неравенств:

$$a_1>0,\quad \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}>0,\quad \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a & a_3 \end{vmatrix}>0,\quad \ldots,\quad \begin{vmatrix} a_1 & a_3 & a_5 \ldots 0 \\ a_0 & a_2 & a_4 \ldots 0 \\ 0 & a_1 & a_3 \ldots 0 \\ 0 & a_0 & a_2 \ldots 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ \cdot & \cdot & \cdot\ \cdot\ \cdot\ \cdot\ a_n \end{vmatrix}>0. \qquad (35')$$

Вещественный многочлен $f(z)=a_0z^n+\ldots$, коэффициенты которого удовлетворяют условиям (34), т. е. вещественный многочлен, у которого все корни имеют отрицательные вещественные части, обычно называют *многочленом Гурвица.*

Отметим два замечательных свойства схемы Рауса:

1. Обозначим элементы $p+1$-й строки схемы Рауса через $\alpha_{p0}, \alpha_{p1}, \alpha_{p2}, \ldots$; тогда $\alpha_{pj}=\dfrac{\Delta_p^{(p+j)}}{\Delta_{p-1}}$ $(p, j=0, 1, \ldots)$. Здесь $\Delta_p^{(p)}=\Delta_p=H\begin{pmatrix}1 \ldots p\\ 1 \ldots p\end{pmatrix}$ — определитель Гурвица, а $\Delta_p^{(p+j)}=H\begin{pmatrix}1 \ldots p-1\ p\\ 1 \ldots p-1\ p+j\end{pmatrix}$ при $j\geqslant 1$ — «побочный» определитель Гурвица p-го порядка. Между элементами схемы Рауса имеет место основная зависимость [см. формулы (13), (13′) на стр. 473]:

$$\alpha_{pj}=\frac{\alpha_{p0}}{\alpha_{p+1,0}}\alpha_{p+1,j}+\alpha_{p+2,j-1} \qquad (p, j=0, 1, \ldots;\ \alpha_{kl}=0, \text{ если } k>n \text{ либо } j<0).$$

Элементы любой p-й строки схемы Рауса получаются из элементов двух последующих строк с помощью двух операций — умножения на отношение $\alpha_{p0}/\alpha_{p+1,0}$ и сложения. Поэтому (в регулярном случае!) элементы произвольной p-й строки схемы Рауса могут быть выражены с помощью операций сложения и умножения через элементы последних двух строк $\alpha_{n-1,0}$ и $\alpha_{n,0}$ и через отношения $\dfrac{\alpha_{p0}}{\alpha_{p+1,0}}, \ldots, \dfrac{\alpha_{n-2,0}}{\alpha_{n-1,0}}$ и представлены в виде

$$\alpha_{pj}=\frac{\varphi_{pj}(\alpha_{p0}, \alpha_{p+1,0}, \ldots, \alpha_{n0})}{\alpha_{p+1,0}\ldots\alpha_{n0}} \qquad (p, j=0, 1, \ldots), \qquad (36)$$

где $\varphi_{pj}(\alpha_{p0}, \alpha_{p+1,0}, \ldots, \alpha_{n0})$ — многочлены с целыми положительными коэффициентами.

С помощью формул (36) все элементы схемы Рауса и, в частности (при $p=0, 1$), коэффициенты исходного многочлена $f(z)$ рационально выражаются (и притом с положительными коэффициентами) через элементы первого столбца схемы Рауса.

Если выполняется критерий Рауса, т. е. все элементы первого столбца схемы Рауса положительны, то из формулы (36) непосредственно следует, что в этом случае все элементы схемы Рауса и, в частности, коэффициенты основного многочлена, положительны.

Заметим еще, что, заменяя в формулах (36) величины α_{pj} через отношения $\Delta_p^{(p+j)}/\Delta_{p-1}$, можно рационально (с положительными коэффициентами) выразить побочные определители Гурвица $\Delta_p^{(p+j)}$ через основные.

2. Пусть $f_0, f_1, \ldots$ и $g_0, g_1, \ldots$ — $(m+1)$-я и $(m+2)$-я строки схемы $\left(f_0=\dfrac{\Delta_m}{\Delta_{m-1}},\ g_0=\dfrac{\Delta_{m+1}}{\Delta_m}\right)$. Так как эти две строки вместе с последующими образуют самостоятельную схему Рауса, то элементы $(m+p+1)$-й строки (в первоначальной схеме) выражаются через элементы $(m+1)$-й и $(m+2)$-й строк, $f_0, f_1, \ldots$ и $g_0, g_1, \ldots$ — по тем же формулам, по каким элементы $(p+1)$-й строки выражаются через элементы первых

двух строк $a_0, a_1, \ldots$ и $b_0, b_1, \ldots$, т. е., полагая

$$\widetilde{H}=\begin{Vmatrix} g_0 & g_1 & g_2 \ldots \\ f_0 & f_1 & f_2 \ldots \\ 0 & g_0 & g_1 \ldots \\ 0 & f_0 & f_1 \ldots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{Vmatrix},$$

будем иметь:

$$\frac{\Delta_{m+p}^{(m+j)}}{\Delta_{m+p-1}}=\frac{\widetilde{\Delta}_p^{(j)}}{\widetilde{\Delta}_{p-1}} \qquad (j=p,\ p+1,\ \ldots). \tag{37}$$

Определитель Гурвица Δ_{m+p} равен произведению первых $m+p$ чисел в ряду $b_0, c_0, \ldots$:

$$\Delta_{m+p}=b_0c_0\ldots f_0g_0\ldots l_0.$$

Но

$$\Delta_m=b_0c_0\ldots f_0, \qquad \widetilde{\Delta}_p=g_0\ldots l_0.$$

Поэтому имеет место следующее важное соотношение:

$$\Delta_{m+p}=\Delta_m\widetilde{\Delta}_p\ ^{1)}. \tag{38}$$

Формула (38) имеет место всегда, если только определены числа $f_0, f_1, \ldots$ и $g_0, g_1, \ldots$, т. е. при условии $\Delta_{m-1}\neq 0$, $\Delta_m\neq 0$.

Формулы (37) имеют смысл, если дополнительно к условиям $\Delta_{m-1}\neq 0$, $\Delta_m\neq 0$ выполняется и условие $\Delta_{m+p-1}\neq 0$. Из этого условия уже следует, что и знаменатель дроби, стоящей в правой части равенства (37), не равен нулю: $\widetilde{\Delta}_{p-1}\neq 0$.

§ 7. Формула Орландо

При рассмотрении случаев, когда некоторые из определителей Гурвица равны нулю, нам понадобится следующая формула Орландо [137], выражающая определитель Δ_{n-1} через старший коэффициент a_0 и корни $z_1, z_2, \ldots, z_n$ многочлена $f(z)$ [2])

$$\Delta_{n-1}=(-1)^{\frac{n(n-1)}{2}}a_0^{n-1}\prod_{i<k}^{1\ldots n}(z_i+z_k). \tag{39}$$

При $n=2$ эта формула сводится к известной формуле для коэффициента b_0 в квадратном уравнении $a_0z^2+b_0z+a_1=0$:

$$\Delta_1=b_0=-a_0(z_1+z_2).$$

Допустим теперь, что формула (39) справедлива для многочлена n-й степени $f(z)=a_0z^n+b_0z^{n-1}+\ldots$, и покажем, что она справедлива для многочлена $(n+1)$-й степени

$$F(z)=(z+h)f(z)=$$
$$=a_0z^{n+1}+(b_0+ha_0)z^n+(a_1+hb_0)z^{n-1}+\ldots \qquad (h=-z_{n+1}).$$

---

[1]) Здесь $\widetilde{\Delta}_p$ — минор p-го порядка, стоящий в левом верхнем углу матрицы $\widetilde{H}$.

[2]) При этом коэффициентами многочлена $f(z)$ могут быть произвольные комплексные числа.

Для этого составим вспомогательный определитель $(n+1)$-го порядка

$$D=\begin{vmatrix} b_0 & b_1 \dots b_{n-1} & h^n \\ a_0 & a_1 \dots a_{n-1} & -h^{n-1} \\ 0 & b_0 \dots b_{n-2} & h^{n-2} \\ 0 & a_0 \dots a_{n-2} & -h^{n-3} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ 0 & 0 \dots\dots & (-1)^n \end{vmatrix} \begin{pmatrix} a_k=0 & \text{при} & k>\left[\frac{n}{2}\right], \\ b_k=0 & \text{при} & k>\left[\frac{n-1}{1}\right] \end{pmatrix}.$$

Помножим первую строку D на a_0 и прибавим к ней вторую, помноженную на $-b_0$, третью, помноженную на a_1, четвертую — на $-b_1$ и т. д. Тогда все элементы первой строки, кроме последнего, обратятся в нуль, а последний элемент будет равен $f(h)$. Отсюда легко заключаем, что

$$D=(-1)^n\Delta_{n-1}f(h).$$

С другой стороны, прибавляя к каждой (кроме последней) строке определителя D последующую, помноженную на h, мы получим помноженный на $(-1)^n$ определитель Гурвица Δ_n^* n-го порядка для многочлена $F(z)$:

$$D=(-1)^n\begin{vmatrix} b_0+ha_0 & b_1+ha_1 \dots \\ a_0 & a_1+hb_0 \dots \\ 0 & b_0+ha_0 \dots \\ 0 & a_0 \dots \\ \dots & \dots \\ \dots & \dots \end{vmatrix}=(-1)^n\Delta_n^*.$$

Таким образом,

$$\Delta_n^*=\Delta_{n-1}f(h)=a_0\Delta_{n-1}\prod_{i=1}^{n}(h-z_i).$$

Заменяя здесь Δ_{n-1} на его выражение из (39) и полагая $h=-z_{n+1}$, получаем:

$$\Delta_n^*=(-1)^{\frac{(n+1)n}{2}}a_0^n\prod_{i<k}^{1\dots n+1}(z_i+z_k).$$

Таким образом, методом математической индукции установлена справедливость формулы Орландо для многочлена любой степени.

Из формулы Орландо следует, что $\Delta_{n-1}=0$ *тогда и только тогда, когда сумма двух корней многочлена* $f(z)$ *равна нулю*[1]).

Так как $\Delta_n=c\Delta_{n-1}$, где c — свободный член многочлена $f(z)$ $(c=(-1)^na_0z_1z_2\dots z_n)$, то из (39) следует:

$$\Delta_n=(-1)^{\frac{n(n+1)}{2}}a_0^nz_1z_2\dots z_n\prod_{i<k}^{1\dots n}(z_i+z_k). \tag{40}$$

Последняя формула показывает, что Δ_n *обращается в нуль тогда и только тогда, когда у* $f(z)$ *существует такой корень* z, *что и* $-z$ *является корнем.*

---

[1]) В частности, $\Delta_{n-1}=0$, когда $f(z)$ имеет хотя бы одну пару сопряженных чисто мнимых корней или кратный нулевой корень.

§ 8. Особые случаи в теореме Рауса — Гурвица

При рассмотрении особых случаев, когда некоторые из определителей Гурвица равны нулю, мы можем предполагать, что $\Delta_n \neq 0$ (и, следовательно, $\Delta_{n-1} \neq 0$).

Действительно, если $\Delta_n=0$, то, как было выяснено в конце предыдущего параграфа, вещественный многочлен $f(z)$ имеет такой корень z', для которого $-z'$ также является корнем $f(z)$. Если положить $f(z)=F_1(z)+F_2(z)$, где

$$F_1(z)=a_0z^n+a_1z^{n-1}+\ldots, \quad F_2(z)=b_0z^{n-1}+b_1z^{n-1}+\ldots,$$

то из равенств $f(z')=f(-z')=0$ можно заключить, что $F_1(z')=F_2(z')=0$. Следовательно, z' будет корнем наибольшего общего делителя $d(z)$ многочленов $F_1(z)$ и $F_2(z)$. Полагая $f(z)=d(z)f^*(z)$, мы сведем задачу Рауса—Гурвица для $f(z)$ к такой же задаче для многочлена $f^*(z)$, для которого уже последний определитель Гурвица отличен от нуля[1]).

1. Рассмотрим сначала тот случай, когда

$$\Delta_1=\ldots=\Delta_p=0, \quad \Delta_{p+1}\neq 0, \ \ldots, \ \Delta_n\neq 0. \tag{41}$$

Из $\Delta_1=0$ следует: $b_0=0$; из $\Delta_2=\begin{vmatrix} 0 & b_1 \\ a_0 & a_1 \end{vmatrix}=-a_0b_1=0$ вытекает: $b_1=0$. Но тогда автоматически

$$\Delta_3=\begin{vmatrix} 0 & b_1 & b_2 \\ a_0 & a_1 & a_2 \\ 0 & 0 & b_1 \end{vmatrix}=-a_0b_1^2=0.$$

Из

$$\Delta_4=\begin{vmatrix} 0 & 0 & b_2 & b_3 \\ a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & b_2 \\ 0 & a_0 & a_1 & a_2 \end{vmatrix}=-a_0^2b_2^2=0$$

следует: $b_2=0$, а тогда $\Delta_5=-a_0^2b_2^3=0$ и т. д.

Приведенные рассуждения показывают, что в (41) всегда p — *нечетное* число: $p=2h-1$. При этом $b_0=b_1=b_2=\ldots=b_{h-1}=0$, $b_h\neq 0$ и [2])

$$\Delta_{p+1}=\Delta_{2h}=(-1)^{\frac{h(h+1)}{2}}a_0^hb_h^h, \quad \Delta_{p+2}=\Delta_{2h+1}=(-1)^{\frac{h(h+1)}{2}}a_0^hb_h^{h+1}=\Delta_{p+1}b_h. \tag{42}$$

---

[1]) В случае, когда

$$\Delta_n=\Delta_{n-1}=\ldots=\Delta_{m+1}=0, \quad \Delta_m\neq 0, \quad \Delta_{m-1}\neq 0, \ \ldots, \ \Delta_1\neq 0,$$

можно в явном виде записать уравнения $d(z)=0$. Действительно, функции $F_1(z)$ и $F_2(z)$ связаны с функциями $f_1(\omega)$ и $f_2(\omega)$ [см. формулу (9) на стр. 472] соотношениями

$$F_1(z)=i^nf_1(-iz), \qquad F_2(z)=i^{n-1}f_2(-iz).$$

Поэтому уравнение $d(z)=0$ совпадает с уравнением $f_{m+1}(-iz)=0$, где многочлен $f_{m+1}(\omega)$ — наибольший общий делитель многочленов $f_1(z)$ и $f_2(z)$ — определяется последней строкой схемы Рауса. Следовательно, в силу формул (31) уравнение $d(z)=0$ может быть записано в виде

$$\sum_{j=0}^{\left[\frac{n-m}{2}\right]}\Delta_m^{(m+j)}z^{n-m-2j}=0,$$

где $\Delta_m^{(m+j)}=H\begin{pmatrix} 1 \ldots m-1 & m \\ 1 \ldots m-1 & m+j \end{pmatrix}$ $(j=0, 1, \ldots)$ — побочные определители Гурвица.

[2]) Из (42) следует, что при h нечетном $\operatorname{sign}\Delta_{p+2}=(-1)^{\frac{h+1}{2}}\operatorname{sign}a_0$, а при h четном $\operatorname{sign}\Delta_{p+1}=(-1)^{\frac{h}{2}}$.

Проварьируем, т. е. изменим немного, коэффициенты $b_0, b_1, \ldots, b_{h-1}$ так, чтобы при новых проварьированных значениях $b_0^*, b_1^*, \ldots, b_{h-1}^*$ все определители Гурвица $\Delta_1^*, \Delta_2^*, \ldots, \Delta_n^*$ стали отличными от нуля и чтобы при этом определители $\Delta_{p+1}^*, \ldots, \Delta_n^*$ сохранили свои прежние знаки. Мы будем считать $b_0^*, b_1^*, \ldots, b_{h+1}^*$ «малыми» величинами разного порядка «малости», а именно мы примем, что каждое b_{j-1}^* по абсолютной величине «значительно» меньше b_j^* $(j=1, 2, \ldots, h;\ b_h^*=b_h)$. Последнее означает, что при вычислении знака целого алгебраического выражения относительно b_i^* мы можем пренебрегать членами, в которых некоторые b_i^* имеют индекс $<j$ по сравнению с членами, где все b_i^* имеют индекс $\geqslant j$. После этого мы легко найдем «знакоопределяющие» члены в $\Delta_1^*, \Delta_2^*, \ldots, \Delta_p^*$ $(p=2h-1)$ [1]):

$$\Delta_1^*=b_0^*,\ \Delta_2^*=-a_0b_1^*+\ldots,\ \Delta_3^*=-a_0b_1^{*2}+\ldots,\ \Delta_4^*=-a_0^2b_2^{*2}+\ldots,$$

$$\Delta_5^*=-a_0^2b_2^{*3}+\ldots,\ \Delta_6^*=a_0^3b_3^{*3}+\ldots$$

и т. д.; вообще

$$\left.\begin{aligned} \Delta_{2j}^*&=(-1)^{\frac{j(j+1)}{2}}a_0^jb_j^{*j}+\ldots && (j=1, 2, \ldots, h-1),\\ \Delta_{2j+1}^*&=(-1)^{\frac{j(j+1)}{2}}a_0^jb_j^{*j+1}+\ldots && (j=0, 1, \ldots, h-1). \end{aligned}\right\} \tag{43}$$

Выберем $b_0^*, b_1^*, \ldots, b_{2h-1}^*$ положительными; тогда знаки Δ_i^* определятся из формулы

$$\operatorname{sign}\Delta_i^*=(-1)^{\frac{j(j+1)}{2}}\operatorname{sign}a_0^j \qquad \left(j=\left[\frac{i}{2}\right],\ i=1, 2, \ldots, p\right). \tag{44}$$

При любом малом варьировании коэффициентов многочлена число k остается неизменным, поскольку многочлен $f(z)$ не имеет корней на мнимой оси. Поэтому, исходя из (44), мы определяем число корней в правой полуплоскости по формуле

$$k=V\left(a_0,\ \Delta_1^*,\ \frac{\Delta_2^*}{\Delta_2^*},\ \ldots,\ \frac{\Delta_{p+1}}{\Delta_p^*},\ \frac{\Delta_{p+2}}{\Delta_{p+1}}\right)+V\left(\frac{\Delta_{p+2}}{\Delta_{p+1}},\ \ldots,\ \frac{\Delta_n}{\Delta_{n-1}}\right). \tag{45}$$

Элементарный подсчет, проведенный на основании формул (42) и (44), показывает, что

$$V\left(a_0,\ \Delta_1^*,\ \frac{\Delta_2^*}{\Delta_1^*},\ \ldots,\ \frac{\Delta_{p+1}}{\Delta_p^*},\ \frac{\Delta_{p+2}}{\Delta_{p+1}}\right)=h+\frac{1-(-1)^h\varepsilon}{2}\left(\begin{matrix} p=2h-1,\\ \varepsilon=\operatorname{sign}\left(a_0\dfrac{\Delta_{p+2}}{\Delta_{p+1}}\right)\end{matrix}\right). \tag{46}$$

Заметим, что величина, стоящая в левой части равенства (46), не зависит от способа варьирования коэффициентов и при любых малых варьированиях сохраняет одно и то же значение. Это следует из формулы (45), поскольку k не меняет своего значения при малом варьировании коэффициентов.

2. Пусть теперь при $s>0$

$$\Delta_{s+1}=\ldots=\Delta_{s+p}=0, \tag{47}$$

а все остальные определители Гурвица отличны от нуля.

Обозначим через $\tilde{a}_0, \tilde{a}_1, \ldots$ и $\tilde{b}_0, \tilde{b}_1, \ldots$ элементы $(s+1)$-й и $(s+2)$-й строки в схеме Рауса $\left(\tilde{a}_0=\dfrac{\Delta_s}{\Delta_{s-1}},\ \tilde{b}_0=\dfrac{\Delta_{s+1}}{\Delta_s}\right)$. Соответствующие определители Гурвица

---

[1]) По существу аналогичные члены уже были вычислены выше для $\Delta_1, \Delta_2, \ldots, \Delta_p$.

§ 8. Особые случаи в теореме Рауса — Гурвица

При рассмотрении особых случаев, когда некоторые из определителей Гурвица равны нулю, мы можем предполагать, что $\Delta_n \neq 0$ (и, следовательно, $\Delta_{n-1} \neq 0$).

Действительно, если $\Delta_n = 0$, то, как было выяснено в конце предыдущего параграфа, вещественный многочлен $f(z)$ имеет такой корень z', для которого $-z'$ также является корнем $f(z)$. Если положить $f(z) = F_1(z) + F_2(z)$, где

$$F_1(z) = a_0 z^n + a_1 z^{n-1} + \ldots, \quad F_2(z) = b_0 z^{n-1} + b_1 z^{n-1} + \ldots,$$

то из равенств $f(z') = f(-z') = 0$ можно заключить, что $F_1(z') = F_2(z') = 0$. Следовательно, z' будет корнем наибольшего общего делителя $d(z)$ многочленов $F_1(z)$ и $F_2(z)$. Полагая $f(z) = d(z) f^*(z)$, мы сведем задачу Рауса—Гурвица для $f(z)$ к такой же задаче для многочлена $f^*(z)$, для которого уже последний определитель Гурвица отличен от нуля[1]).

1. Рассмотрим сначала тот случай, когда

$$\Delta_1 = \ldots = \Delta_p = 0, \quad \Delta_{p+1} \neq 0, \ \ldots, \ \Delta_n \neq 0. \tag{41}$$

Из $\Delta_1 = 0$ следует: $b_0 = 0$; из $\Delta_2 = \begin{vmatrix} 0 & b_1 \\ a_0 & a_1 \end{vmatrix} = -a_0 b_1 = 0$ вытекает: $b_1 = 0$. Но тогда автоматически

$$\Delta_3 = \begin{vmatrix} 0 & b_1 & b_2 \\ a_0 & a_1 & a_2 \\ 0 & 0 & b_1 \end{vmatrix} = -a_0 b_1^2 = 0.$$

Из

$$\Delta_4 = \begin{vmatrix} 0 & 0 & b_2 & b_3 \\ a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 0 & b_2 \\ 0 & a_0 & a_1 & a_2 \end{vmatrix} = -a_0^2 b_2^2 = 0$$

следует: $b_2 = 0$, а тогда $\Delta_5 = -a_0^2 b_2^2 = 0$ и т. д.

Приведенные рассуждения показывают, что в (41) всегда p — *нечетное* число: $p = 2h - 1$. При этом $b_0 = b_1 = b_2 = \ldots = b_{h-1} = 0$, $b_h \neq 0$ и [2])

$$\Delta_{p+1} = \Delta_{2h} = (-1)^{\frac{h(h+1)}{2}} a_0^h b_h^h, \quad \Delta_{p+2} = \Delta_{2h+1} = (-1)^{\frac{h(h+1)}{2}} a_0^h b_h^{h+1} = \Delta_{p+1} b_h. \tag{42}$$

---

[1]) В случае, когда

$$\Delta_n = \Delta_{n-1} = \ldots = \Delta_{m+1} = 0, \quad \Delta_m \neq 0, \quad \Delta_{m-1} \neq 0, \ \ldots, \ \Delta_1 \neq 0,$$

можно в явном виде записать уравнения $d(z) = 0$. Действительно, функции $F_1(z)$ и $F_2(z)$ связаны с функциями $f_1(\omega)$ и $f_2(\omega)$ [см. формулу (9) на стр. 472] соотношениями

$$F_1(z) = i^n f_1(-iz), \qquad F_2(z) = i^{n-1} f_2(-iz).$$

Поэтому уравнение $d(z) = 0$ совпадает с уравнением $f_{m+1}(-iz) = 0$, где многочлен $f_{m+1}(\omega)$ — наибольший общий делитель многочленов $f_1(z)$ и $f_2(z)$ — определяется последней строкой схемы Рауса. Следовательно, в силу формул (31) уравнение $d(z) = 0$ может быть записано в виде

$$\sum_{j=0}^{\left[\frac{n-m}{2}\right]} \Delta_m^{(m+j)} z^{n-m-2j} = 0,$$

где $\Delta_m^{(m+j)} = H\begin{pmatrix} 1 \ldots m-1 & m \\ 1 \ldots m-1 & m+j \end{pmatrix}$ $(j = 0, 1, \ldots)$ — побочные определители Гурвица.

[2]) Из (42) следует, что при h нечетном $\operatorname{sign} \Delta_{p+2} = (-1)^{\frac{h+1}{2}} \operatorname{sign} a_0$, а при h четном $\operatorname{sign} \Delta_{p+1} = (-1)^{\frac{h}{2}}$.

в которой при подсчете величины V для каждой группы подряд идущих p нулевых определителей (p — всегда нечетное число!)

$$(\Delta_s \neq 0) \quad \Delta_{s+1} = \ldots = \Delta_{s+p} = 0 \quad (\Delta_{s+p+1} \neq 0)$$

следует положить:

$$V\left(\frac{\Delta_s}{\Delta_{s-1}}, \frac{\Delta_{s+1}}{\Delta_s}, \ldots, \frac{\Delta_{s+p+2}}{\Delta_{s+p+1}}\right) = h + \frac{1-(-1)^h \varepsilon}{2}, \tag{50}$$

где

$$p = 2h-1 \quad \text{и} \quad \varepsilon = \operatorname{sign}\left(\frac{\Delta_s}{\Delta_{s-1}} \frac{\Delta_{s+p+2}}{\Delta_{s+p+1}}\right)^{1)}.$$

§ 9. Метод квадратичных форм. Определение числа различных вещественных корней многочлена

Раус получил свой алгоритм, применяя теорему Штурма к вычислению индекса Коши правильной рациональной дроби специального типа [см. формулу (11) на стр. 473]. У этой дроби из двух многочленов — числителя и знаменателя — один содержит только четные, а другой только нечетные степени аргумента z.

В настоящем параграфе и в последующих параграфах мы изложим более глубокий и более перспективный метод квадратичных форм Эрмита в применении к проблеме Рауса — Гурвица. При помощи этого метода мы получим выражение для индекса произвольной рациональной дроби через коэффициенты числителя и знаменателя. Метод квадратичных форм позволяет применить к проблеме Рауса — Гурвица результаты тонких исследований Фробениуса по теории ганкелевых форм (гл. X, § 10) и установить тесную связь некоторых замечательных теорем П. Л. Чебышева и А. А. Маркова с задачей устойчивости.

Мы познакомим читателя с методом квадратичных форм сначала на сравнительно простой задаче определения числа различных вещественных корней многочлена.

При решении этой задачи мы можем ограничиться случаем, когда $f(z)$ — вещественный многочлен. Действительно, пусть дан комплексный многочлен $f(z) = u(z) + iv(z)$ [$u(z)$ и $v(z)$ — вещественные многочлены]. Каждый вещественный корень многочлена $f(z)$ обращает в нуль одновременно и $u(z)$ и $v(z)$. Поэтому комплексный многочлен $f(z)$ имеет те же вещественные корни, что и вещественный многочлен $d(z)$, являющийся наибольшим общим делителем многочленов $u(z)$ и $v(z)$.

Итак, пусть $f(z)$ — вещественный многочлен с различными корнями $\alpha_1, \alpha_2, \ldots, \alpha_q$ соответственно кратностей $n_1, n_2, \ldots, n_q$:

$$f(z) = a_0 (z-\alpha_1)^{n_1} (z-\alpha_2)^{n_2} \ldots$$
$$\ldots (z-\alpha_q)^{n_q} \, [a_0 \neq 0; \quad \alpha_i \neq \alpha_k \quad \text{при } i \neq k \quad (i, k = 1, 2, \ldots, q)].$$

Введем в рассмотрение суммы Ньютона

$$s_p = \sum_{j=1}^{q} n_j \alpha_j^p \qquad (p = 0, 1, 2, \ldots).$$

---

1) При $s=1$ отношение $\frac{\Delta_s}{\Delta_{s-1}}$ следует заменить на Δ_1, а при $s=0$ — на a_0.

При помощи этих сумм составим ганкелеву форму

$$S_n(x, x) = \sum_{i, k=0}^{n-1} s_{i+k} x_i x_k,$$

где n — любое целое число $\geqslant q$.

Тогда имеет место следующая

Теорема 6. *Число всех различных корней многочлена $f(z)$ равно рангу, а число всех различных вещественных корней равно сигнатуре формы $S_n(x, x)$.*

Доказательство. Из определения формы $S_n(x, x)$ непосредственно вытекает следующее ее представление:

$$S_n(x, x) = \sum_{j=1}^{q} n_j (x_0 + \alpha_j x_1 + \alpha_j^2 x_2 + \ldots + \alpha_j^{n-1} x_{n-1})^2. \tag{51}$$

Здесь каждому корню α_j многочлена $f(z)$ соответствует квадрат линейной формы $Z_j = x_0 + \alpha_j x_1 + \ldots + \alpha_j^{n-1} x_{n-1}$ $(j = 1, 2, \ldots, q)$. Формы Z_1, $Z_2, \ldots, Z_q$ линейно независимы между собой, так как коэффициенты этих линейных форм образуют матрицу Вандермонда $\| \alpha_j^h \|$, ранг которой равен числу различных α_j, т. е. q. Следовательно (см. стр. 269), ранг r формы $S_n(x, x)$ равен q.

В представлении (51) каждому вещественному корню α_j отвечает положительный квадрат. Каждой паре комплексно сопряженных корней α_j и $\bar{\alpha}_j$ отвечают две комплексно сопряженные формы:

$$Z_j = P_j + iQ_j, \qquad \bar{Z}_j = P_j - iQ_j;$$

соответствующие слагаемые в (51) в сумме дают один положительный и один отрицательный квадрат:

$$n_j Z_j^2 + n_j \bar{Z}_j^2 = 2n_j P_j^2 - 2n_j Q_j^2.$$

Отсюда легко усмотреть[1]), что сигнатура формы $S_n(x, x)$, т. е. разность между числом положительных и числом отрицательных квадратов, равна числу различных вещественных α_j.

Теорема доказана.

Из доказанной теоремы вытекает, что *все формы*

$$S_n(x, x) \qquad (n = q, q+1, \ldots)$$

имеет один и тот же ранг и одну и ту же сигнатуру.

Применяя теорему 6 к определению числа различных вещественных корней, возьмем в качестве n степень многочлена $f(z)$. Используя установленное в главе X (стр. 275) правило определения сигнатуры квадратичной формы, мы получаем

Следствие. *Число различных вещественных корней вещественного многочлена $f(z)$ равно избытку числа постоянств знака над числом перемен знака в ряду чисел*

$$1, \; s_0, \; \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix}, \; \ldots, \; \begin{vmatrix} s_0 & s_1 & \ldots & s_{r-1} \\ s_1 & s_2 & \ldots & s_r \\ \cdot & \cdot & \cdot & \cdot \\ s_{r-1} & s_r & \ldots & s_{2r-2} \end{vmatrix}, \tag{52}$$

[1]) Квадратичная форма $S_n(x, x)$ представлена в виде суммы (алгебраической) q квадратов вещественных форм Z_j (для вещественных λ_j), P_j и Q_j (для комплексных λ_j). Эти формы независимы, так как число q равняется рангу r формы $S_n(x, x)$.

где s_p $(p=0, 1, 2, \ldots)$ — *суммы Ньютона для многочлена* $f(z)$, *а* r — ранг ганкелевой формы $S_n(x, x)=\sum\limits_{i,\,k=0}^{n-1} s_{i+k}x_ix_k$ [n — *степень многочлена* $f(z)$].

Сформулированное таким образом правило для определения числа различных вещественных корней непосредственно применимо лишь в случае, когда все числа в ряду (52) отличны от нуля. Однако, поскольку здесь идет речь о вычислении сигнатуры ганкелевой квадратичной формы, то на основе результатов главы X, § 10 это правило с надлежащими уточнениями применяется в самом общем случае (более подробно об этом см. § 11 этой главы).

Число различных вещественных корней вещественного многочлена $f(z)$ равно индексу $I_{-\infty}^{+\infty}\dfrac{f'(z)}{f(z)}$ (см. стр. 470). Поэтому следствие из теоремы 6 дает нам формулу

$$I_{-\infty}^{+\infty}\frac{f'(z)}{f(z)}=r-2V\left(1,\ s_0,\ \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix},\ \ldots,\ \begin{vmatrix} s_0 & s_1 & \ldots & s_{r-1} \\ s_1 & s_2 & \ldots & s_r \\ \cdot & \cdot & \cdot & \cdot \\ s_{r-1} & s_r & \ldots & s_{2r-2} \end{vmatrix}\right).$$

В § 11 мы установим аналогичную формулу для индекса произвольной рациональной дроби. Необходимые для этого сведения о бесконечных ганкелевых матрицах будут даны в следующем параграфе.

§ 10. Бесконечные ганкелевы матрицы конечного ранга

1. Пусть дана последовательность комплексных чисел

$$s_0,\ s_1,\ s_2,\ \ldots$$

Эта последовательность чисел определяет бесконечную симметрическую матрицу

$$S=\begin{Vmatrix} s_0 & s_1 & s_2 & \ldots \\ s_1 & s_2 & s_3 & \ldots \\ s_2 & s_3 & s_4 & \ldots \\ \cdot & \cdot & \cdot & \cdot \end{Vmatrix},$$

которую называют обычно *ганкелевой*. Наряду с бесконечными ганкелевыми матрицами рассматриваются конечные ганкелевы матрицы $S_n=\|s_{i+k}\|_0^{n-1}$ и связанные с ними ганкелевы формы

$$S_n(x, x)=\sum_{i,\,k=0}^{n-1} s_{i+k}x_ix_k.$$

Последовательные главные миноры матрицы S будем обозначать через $D_1, D_2, D_3, \ldots$:

$$D_p=|s_{i+k}|_0^{p-1} \qquad (p=1, 2, \ldots).$$

Бесконечные матрицы могут быть конечного и бесконечного ранга. В последнем случае в этих матрицах существуют отличные от нуля миноры сколь угодно большого порядка. Следующая теорема дает необходимое и достаточное условие, которому должна удовлетворять последовательность чисел $s_0, s_1, s_2, \ldots$ для того, чтобы порождаемая ею бесконечная ганкелева матрица $S=\|s_{i+k}\|_0^{\infty}$ имела конечный ранг.

Теорема 7. *Бесконечная матрица* $S=\|s_{i+k}\|_0^\infty$ *имеет конечный ранг* r *тогда и только тогда, когда существует* r *чисел* $\alpha_1, \alpha_2, \ldots, \alpha_r$ *таких, что*

$$s_q=\sum_{g=1}^{r}\alpha_g s_{q-g} \qquad (q=r,\ r+1,\ \ldots), \tag{53}$$

и r *есть наименьшее число, обладающее этим свойством.*

Доказательство. Если матрица $S=\|s_{i+k}\|_0^\infty$ имеет конечный ранг r, то первые $r+1$ строк $\Gamma_1, \Gamma_2, \ldots, \Gamma_{r+1}$ этой матрицы линейно зависимы. Поэтому существует число $h\leqslant r$ такое, что строки $\Gamma_1, \Gamma_2, \ldots, \Gamma_h$ линейно назависимы, а строка Γ_{h+1} есть линейная комбинация этих строк:

$$\Gamma_{h+1}=\sum_{g=1}^{h}\alpha_g\Gamma_{h-g+1}.$$

Рассмотрим строки $\Gamma_{q+1}, \Gamma_{q+2}, \ldots, \Gamma_{q+h+1}$, где q — любое целое неотрицательное число. Из структуры матрицы S непосредственно видно, что строки $\Gamma_{q+1}, \Gamma_{q+2}, \ldots, \Gamma_{q+h+1}$ получаются из строк $\Gamma_1, \Gamma_2, \ldots, \Gamma_{h+1}$ «укорочением», отбрасыванием элементов, стоящих в первых q столбцах. Поэтому

$$\Gamma_{q+h+1}=\sum_{g=1}^{h}\alpha_g\Gamma_{q+h-g+1} \qquad (q=0,\ 1,\ 2,\ \ldots).$$

Таким образом, в матрице S любая строка, начиная с $(h+1)$-й, выражается линейно через h предыдущих и, следовательно, выражается линейно через h линейно независимых первых строк. Отсюда следует, что *для матрицы* S ранг $r=h$[1]). Линейная зависимость $\Gamma_{q+h+1}=\sum_{g=1}^{h}\alpha_g\Gamma_{q+h-g+1}$ после замены h на r в более подробной записи дает (53).

Обратно, если выполняется условие (53), то в матрице S любая строка (столбец) является линейной комбинацией первых r строк (столбцов). Поэтому все миноры матрицы S, порядок которых $>r$, равны нулю, и матрица S имеет конечный ранг $\leqslant r$. Но этот ранг не может быть $<r$, так как тогда, как было уже показано, имели бы место соотношения вида (53) при меньшем значении r, а это противоречит условию 2). Таким образом, теорема доказана полностью.

Следствие. *Если бесконечная ганкелева матрица* $S=\|s_{i+k}\|_0^\infty$ *имеет конечный ранг* r, *то*

$$D_r=|s_{i+k}|_0^{r-1}\neq 0.$$

Действительно, из соотношений (53) следует, что любая строка (столбец) матрицы S есть линейная комбинация первых r строк (столбцов). Поэтому любой минор r-го порядка матрицы S может быть представлен в виде αD_r, где α — некоторое число. Отсюда следует неравенство $D_r\neq 0$.

---

[1]) Положение «число линейно независимых строк в прямоугольной матрице равно рангу этой матрицы» справедливо не только для конечных, но и для бесконечных строк.

Примечание. Для конечных ганкелевых матриц ранга r неравенство $D_r \neq 0$ может не иметь места. Так, например, матрица $S_2 = \begin{Vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{Vmatrix}$ при $s_0 = s_1 = 0$, $s_2 \neq 0$ имеет ранг 1, в то время как $D_1 = s_0 = 0$.

2. Выясним замечательные взаимные связи между бесконечными ганкелевыми матрицами и рациональным функциями.

Пусть дана правильная рациональная дробная функция

$$R(z) = \frac{g(z)}{h(z)},$$

где

$$h(z) = a_0 z^m + \ldots + a_m \ (a_0 \neq 0), \quad g(z) = b_1 z^{m-1} + b_2 z^{m-2} + \ldots + b_m.$$

Напишем разложение $R(z)$ в степенной ряд по отрицательным степеням z:

$$R(z) = \frac{g(z)}{h(z)} = \frac{s_0}{z} + \frac{s_1}{z^2} + \frac{s_2}{z^3} + \ldots$$

Если все полюсы функции $R(z)$, т. е. все значения z, при которых $R(z)$ обращается в бесконечность, лежат в круге $|z| \leqslant a$, то ряд, стоящий в правой части разложения, сходится при $|z| > a$. Обе части последнего равенства помножим на знаменатель $h(z)$:

$$(a_0 z^m + a_1 z^{m-1} + \ldots + a_m)\left(\frac{s_0}{z} + \frac{s_1}{z^2} + \frac{s_2}{z^3} + \ldots\right) = \\ = b_1 z^{m-1} + b_2 z^{m-2} + \ldots + b_m.$$

Приравнивая между собой коэффициенты при одинаковых степенях z в обеих частях этого тождества, получим следующую систему соотношений:

$$\left.\begin{aligned} &a_0 s_0 = b_1, \\ &a_0 s_1 + a_1 s_0 = b_2, \\ &\ldots\ldots\ldots \\ &a_0 s_{m-1} + a_1 s_{m-2} + \ldots + a_{m-1} s_0 = b_m, \end{aligned}\right\} \tag{54}$$

$$a_0 s_q + a_1 s_{q-1} + \ldots + a_m s_{q-m} = 0 \qquad (q = m, m+1, \ldots). \tag{54'}$$

Полагая

$$\alpha_g = -\frac{a_g}{a_0} \qquad (g = 1, 2, \ldots, m),$$

мы можем соотношения (54′) записать в виде (53) (при $r = m$). Следовательно, согласно теореме 7 построенная при помощи коэффициентов s_0, s_1, s_2, ... бесконечная ганкелева матрица

$$S = \| s_{i+k} \|_0^\infty$$

имеет конечный ранг ($\leqslant m$).

Обратно, если матрица $S = \| s_{i+k} \|_0^\infty$ имеет конечный ранг r, то имеют место соотношения (53), которые могут быть переписаны в виде (54′) (при $m = r$). Тогда, определяя числа b_1, b_2, ..., b_m равенствами (54), будем иметь разложение

$$\frac{b_1 z^{m-1} + \ldots + b_m}{a_0 z^m + a_1 z^{m-1} + \ldots + a_m} = \frac{s_0}{z} + \frac{s_1}{z^2} + \ldots \tag{54''}$$

Наименьшая степень знаменателя m, при которой имеет место это разложение, совпадает с наименьшим числом m, при котором имеют место соотношения (53). По теореме 7 это наименьшее значение m равно рангу матрицы $S = \| s_{i+k} \|_0^\infty$.

При этом значении m рациональная дробь, стоящая в левой части равенства (54″), является несократимой.

Таким образом, нами доказана следующая теорема:

Теорема 8. *Матрица $S = \| s_{i+k} \|_0^\infty$ имеет конечный ранг в том и только в том случае, когда сумма ряда*

$$R(z) = \frac{s_0}{z} + \frac{s_1}{z^2} + \frac{s_2}{z^3} + \ldots$$

есть рациональная функция переменной z. В этом случае ранг матрицы S совпадает с числом полюсов функции $R(z)$, считая каждый полюс столько раз, какова его кратность.

§ 11. Определение индекса произвольной рациональной дроби через коэффициенты числителя и знаменателя

1. Пусть дана произвольная рациональная функция. Напишем ее разложение в ряд по нисходящим степеням z [1]):

$$R(z) = s_{-u-1}z^u + \ldots + s_{-2}z + s_{-1} + \frac{s_0}{z} + \frac{s_1}{z^2} + \ldots \tag{55}$$

Последовательность коэффициентов при отрицательных степенях z

$$s_0, \ s_1, \ s_2, \ \ldots$$

определяет бесконечную ганкелеву матрицу $S = \| s_{i+k} \|_0^\infty$.

Таким образом, устанавливается соответствие

$$R(z) \sim S.$$

Очевидно, что двум рациональным функциям, разность между которыми есть целая функция, отвечает одна и та же матрица S. Однако не всякая матрица $S = \| s_{i+k} \|_0^\infty$ соответствует рациональной функции. В предыдущем параграфе было установлено, что матрица S тогда и только тогда соответствует рациональной функции, когда эта бесконечная матрица имеет конечный ранг. Этот ранг равен числу полюсов (с учетом кратностей) функции $R(z)$, т. е. равен степени знаменателя $f(z)$ в несократимой дроби $\frac{g(z)}{f(z)} = R(z)$. С помощью разложения (55) устанавливается взаимно однозначное соответствие между правильными рациональными функциями $R(z)$ и ганкелевыми матрицами $S = \| s_{i+k} \|_0^\infty$ конечного ранга.

Отметим некоторые свойства соответствия:

1° Если $R_1(z) \sim S_1$, $R_2(z) \sim S_2$, то при любых числах c_1, c_2

$$c_1R_1(z) + c_2R_2(z) \sim c_1S_1 + c_2S_2.$$

В дальнейшем нам придется встретиться со случаем, когда коэффициенты числителя и знаменателя $R(z)$ будут целыми рациональными функциями параметра α; тогда и R будет рациональной функцией от z и α. Из разложения (54) следует, что в этом случае и числа $s_0, s_1, s_2, \ldots,$

---

[1]) Ряд (55) сходится вне любого круга (с центром в точке $z = 0$), содержащего все полюсы функции $R(z)$.

т. е. элементы матрицы S, будут рационально зависеть от α. Дифференцируя по α почленно разложение (55), получим:

2° Если $R(z, \alpha) \sim S(\alpha)$, то $\frac{\partial R}{\partial \alpha} \sim \frac{\partial S}{\partial \alpha}$ [1]).

2. Напишем разложение $R(z)$ на простейшие дроби:

$$R(z) = Q(z) + \sum_{i=1}^{q} \left\{ \frac{A_1^{(j)}}{z-\alpha_j} + \frac{A_2^{(j)}}{(z-\alpha_j)^2} + \ldots + \frac{A_{\nu_j}^{(j)}}{(z-\alpha_j)^{\nu_j}} \right\}, \tag{56}$$

где $Q(z)$ — многочлен, и покажем, как по числам α и A построить матрицу S, соответствующую рациональной функции $R(z)$.

Для этого рассмотрим сначала простейшую рациональную дробь

$$\frac{1}{z-\alpha} = \sum_{p=0}^{\infty} \frac{\alpha^p}{z^{p+1}}.$$

Ей отвечает матрица

$$S_\alpha = \| \alpha^{i+k} \|_0^\infty.$$

Соответствующая этой матрице форма $S_{\alpha n}(x, x)$ имеет вид

$$S_{\alpha n}(x, x) = \sum_{i, k=0}^{n-1} \alpha^{i+k} x_i x_k = (x_0 + \alpha x_1 + \ldots + \alpha^{n-1} x_{n-1})^2.$$

Если

$$R(z) = Q(z) + \sum_{j=1}^{q} \frac{A^{(j)}}{z-\alpha_j},$$

то в силу 1° соответствующая матрица S определится по формуле

$$S = \sum_{j=1}^{q} A^{(j)} S_{\alpha_j} = \| \sum_{j=1}^{q} A^{(j)} \alpha_j^{i+k} \|_0^\infty,$$

а соответствующие квадратичные формы имеют вид

$$S_n(x, x) = \sum_{j=1}^{q} A^{(j)} (x_0 + \alpha_j x_1 + \ldots + \alpha_j^{n-1} x_{n-1})^2.$$

Для того чтобы перейти к общему случаю (56), мы предварительно $h-1$ раз продифференцируем почленно соотношение

$$\frac{1}{z-\alpha} \sim S_\alpha = \| \alpha^{i+k} \|_0^\infty.$$

Получим согласно 1° и 2° [2]):

$$\frac{1}{(z-\alpha)^h} \sim \frac{1}{(h-1)!} \frac{\partial^{h-1} S_\alpha}{\partial \alpha^{h-1}} = \| C_{i+k}^{h-1} \alpha^{i+k-h+1} \|_0^\infty \quad (C_{i+k}^{h-1} = 0 \text{ при } i+k < h-1).$$

Поэтому, пользуясь снова правилом 1°, в общем случае, когда для $R(z)$ имеет место разложение (56), находим:

$$R(z) \sim S = \sum_{j=1}^{q} \left(A_1^{(j)} + A_2^{(j)} \frac{\partial}{\partial \alpha_j} + \ldots + \frac{1}{(\nu_j - 1)!} A_{\nu_j}^{(j)} \frac{\partial^{\nu_j - 1}}{\partial \alpha_j^{\nu_j - 1}} \right) S_{\alpha_j}. \tag{57}$$

---

[1]) Если $S = \| s_{i+k} \|_0^\infty$, то $\frac{\partial S}{\partial \alpha} = \left\| \frac{\partial s_{i+k}}{\partial \alpha} \right\|_0^\infty$.

[2]) Здесь и ниже C_d^h означает число сочетаний из d по h.

Выполняя дифференцирование, получим:

$$S = \left\| \sum_{j=1}^{q} A_1^{(j)} \alpha_j^{i+k} + A_2^{(j)} C_{i+k}^1 \alpha_j^{i+k-1} + \ldots + A_{\nu_j}^{(j)} C_{i+k}^{\nu_j-1} \alpha_j^{i+k-\nu_j+1} \right\|_0^\infty. \quad (57')$$

Соответствующая ганкелева форма $S_n(x, x) = \sum_{i,k=0}^{n-1} s_{i+k} x_i x_k$ будет равна

$$S_n(x, x) = \sum_{j=1}^{q} \left(A_1^{(j)} + A_2^{(j)} \frac{\partial}{\partial \alpha_j} + \ldots + \frac{1}{(\nu_j - 1)!} A_{\nu_j}^{(j)} \frac{\partial^{\nu_j - 1}}{\partial \alpha_j^{\nu_j - 1}} \right) \times$$
$$\times (x_0 + \alpha_j x_1 + \ldots + \alpha_j^{n-1} x_{n-1})^2. \quad (57'')$$

3. Теперь мы имеем возможность сформулировать и доказать основную теорему [1]):

Теорема 9. *Если*

$$R(z) \sim S$$

и m — ранг матрицы S [2]), *то индекс Коши $I_{-\infty}^{+\infty} R(z)$ равен сигнатуре* [3]) *формы $S_n(x, x)$ при любом $n \geqslant m$:*

$$I_{-\infty}^{+\infty} R(z) = \sigma [S_n(x, x)].$$

Доказательство. Пусть имеет место разложение (56). Тогда согласно (57)

$$S = \sum_{j=1}^{q} T_{\alpha_j},$$

где каждое слагаемое имеет вид

$$T_\alpha = \left(A_1 + A_2 \frac{\partial}{\partial \alpha} + \ldots + \frac{1}{(\nu - 1)!} A_\nu \frac{\partial^{\nu - 1}}{\partial \alpha^{\nu - 1}} \right) S_\alpha, \quad S_\alpha = \| \alpha^{i+k} \|_0^\infty \quad (58)$$

и

$$S_n(x, x) = \sum_{j=1}^{q} T_{\alpha_j}(x, x) = \sum_{\alpha_j \text{вещ.}} T_{\alpha_j}(x, x) + \sum_{\alpha_j \text{компл.}} [T_{\alpha_j}(x, x) + T_{\bar{\alpha}_j}(x, x)].$$

Согласно теореме 8 ранг матрицы T_{α_j} и, следовательно, ранг формы $T_{\alpha_j}(x, x)$ равен ν_j $(j = 1, 2, \ldots, q)$, а ранг $S_n(x, x)$ равен $m = \sum_{j=1}^{q} \nu_j$. Но если ранг суммы нескольких вещественных квадратичных форм равен сумме рангов слагаемых форм, то такое же соотношение имеет место и для сигнатур:

$$\sigma [S_n(x, x)] = \sum_{\alpha_j \text{вещ.}} \sigma [T_{\alpha_j}(x, x)] + \sum_{\alpha_j \text{компл.}} \sigma [T_{\alpha_j}(x, x) + T_{\bar{\alpha}_j}(x, x)]. \quad (59)$$

Рассмотрим раздельно два случая:

1) α вещественно. При любой вариации параметров $A_1, A_2, \ldots, A_{\nu-1}$ и α в

$$\frac{A_1}{z - \alpha} + \frac{A_2}{(z - \alpha)^2} + \ldots + \frac{A_\nu}{(z - \alpha)^\nu} \qquad (A_\nu \neq 0) \quad (60)$$

---

[1]) Эта теорема была доказана Эрмитом в 1856 г. для простейшего случая, когда $R(z)$ не имеет кратных полюсов [195]. В общем случае эта теорема была доказана Гурвицем [195] (см. также [16], стр. 17—19). Приведенное в тексте доказательство отличается от доказательства Гурвица.

[2]) Как мы уже отмечали, m равно степени знаменателя в несократимом представлении рациональной дроби $R(z)$ (см. теорему 8 на стр. 498).

[3]) Сигнатуру формы $S_n(x, x)$ будем обозначать через $\sigma [S_n(x, x)]$.