ROZDZIAŁ XVI

OBLICZANIE PIERWIASTKÓW RÓWNAŃ ALGEBRAICZNYCH

§ 1. Twierdzenie Sturma. Niech $f(x)$ oznacza wielomian o współczynnikach rzeczywistych, nie posiadający pierwiastków wielokrotnych. Wielomiany $f(x)$ i $f_1(x)=f'(x)$ są więc w myśl twierdzenia 12, Rozdziału VIII (§ 9, str. 120) względnie pierwsze. Zastosujmy względem nich algorytm Euklidesa z tą jedynie różnicą znakowania, że przy każdym kolejnym dzieleniu zmienimy znak reszty. Będzie więc:

$$
(1)\qquad
\begin{aligned}
f(x) &= q_1(x)f_1(x)-f_2(x),\\
f_1(x) &= q_2(x)f_2(x)-f_3(x),\\
&\ldots\ldots\ldots\ldots\\
f_{k-2}(x) &= q_{k-1}(x)f_{k-1}(x)-f_k(x),\\
f_{k-1}(x) &= q_k(x)f_k(x),
\end{aligned}
$$

przy czym $f_k(x)$ będzie stałą różną od zera, gdyż wielomiany $f(x)$ i $f_1(x)$ są względnie pierwsze.

Weźmy pod uwagę dla danej wartości rzeczywistej x ciąg liczb:

$$
(2)\qquad f_0(x)=f(x),\quad f_1(x),\quad f_2(x),\quad \ldots,\quad f_k(x),
$$

zwany *ciągiem Sturma*.

Jeżeli w tym ciągu, po ewentualnym odrzuceniu wyrazów pośrednich równych zeru, następuje po dodatnim wyrazie ujemny lub na odwrót, to będziemy mówili, że w ciągu występuje (co najmniej jedna) *zmiana znaku*.

Dla każdej wartości liczby rzeczywistej x możemy obliczyć liczbę $Z(x)$ kolejnych zmian znaku w ciągu (2).

PRZYKŁADY. 1. Niech $f(x)=x^3-x$.

Mamy tu $f_1(x)=f'(x)=3x^2-1$, dalej $x^3-x=\frac{x}{3}(3x^2-1)-\frac{2}{3}x$, zatem $f_2(x)=\frac{2}{3}x$, wreszcie $3x^2-1=\frac{9}{2}x\,\frac{2}{3}x-1$, skąd $f_3=1$.

Dla $x=-2$ ciąg (2) jest ciągiem liczb $-6,\ 11,\ -\frac{4}{3},\ 1$; mamy w nim 3 zmiany znaku czyli $Z(-2)=3$. Dla $x=-1$ ciąg (2) jest ciągiem liczb $0,\ 2,\ -\frac{2}{3},\ 1$; mamy w nim 2 zmiany znaku, zatem $Z(-1)=2$.

Dla $x=0$ ciąg (2) jest ciągiem liczb $0,\ -1,\ 0,\ 1$; mamy w nim 1 zmianę znaku, zatem $Z(0)=1$.

Dla $x=1$ ciąg (2) jest ciągiem liczb $0,\ 2,\ \frac{2}{3},\ 1$; nie ma tu żadnej zmiany znaku, zatem $Z(1)=0$.

2. Niech $f(x)=x^5-3x^3+2x$.

Mamy tu $f'(x)=5x^4-9x^2+2$, dalej:

$$x^5-3x^3+2x=\frac{1}{5}x(5x^4-9x^2+2)-\left(\frac{6}{5}x^3-\frac{8}{5}x\right),$$

$$5x^4-9x^2+2=\frac{25}{6}x\left(\frac{6}{5}x^3-\frac{8}{5}x\right)-\left(\frac{7}{3}x^2-2\right),$$

$$\frac{6}{5}x^3-\frac{8}{5}x=\frac{18}{35}x\left(\frac{7}{3}x^2-2\right)-\frac{20}{35}x,$$

$$\frac{7}{3}x^2-2=\frac{245}{60}x\cdot\frac{20}{35}x-2,$$

zatem:

$$f_0(x)=x^5-3x^3+2x,\qquad f_1(x)=5x^4-9x^2+2,\qquad f_2(x)=\frac{6}{5}x^3-\frac{8}{5}x,$$

$$f_3(x)=\frac{7}{3}x^2-2,\qquad f_4(x)=\frac{20}{35}x,\qquad f_5(x)=2.$$

Dla $x=-2$ ciąg (2) jest ciągiem liczb $-12,\ 46,\ -\frac{32}{5},\ \frac{22}{3},\ -\frac{40}{35},\ 2$, zatem $Z(-2)=5$.

Dla $x=-1$ ciąg (2) jest ciągiem liczb $0,\ -2,\ \frac{2}{5},\ \frac{1}{3},\ -\frac{20}{35},\ 2$, zatem $Z(-1)=3$.

Dla $x=0$ ciąg (2) jest ciągiem liczb $0,\ 2,\ 0,\ -2,\ 0,\ 2$, zatem $Z(0)=2$.

Dla $x=1$ ciąg (2) jest ciągiem liczb $0,\ -2,\ -\frac{2}{5},\ \frac{1}{3},\ \frac{20}{35},\ 2$, zatem $Z(1)=1$.

Wreszcie, dla $x=2$ ciąg (2) jest ciągiem liczb $12,\ 46,\ \frac{32}{5},\ \frac{22}{3},\ \frac{40}{35},\ 2$, zatem $Z(2)=0$.

Zbadamy obecnie dla jakich wartości x zachodzi zmiana wartości funkcji $Z(x)$.

Ponieważ funkcje (2), jako wielomiany, są wszystkie ciągłe, więc jeżeli dla danej wartości $x=x_0$ żadna z liczb (2) nie jest zerem, to dla liczb x dostatecznie bliskich liczby x_0 znaki liczb (2) będą wszystkie odpowiednio te same co dla liczby x_0, a przeto dla x dostatecznie bliskich liczby x_0 będzie $Z(x)=Z(x_0)$. Zmiana wartości liczby $Z(x)$ może więc nastąpić jedynie wówczas, gdy przynajmniej jedna z liczb (2) jest zerem, czyli gdy x przechodzi przez pierwiastek którejkolwiek z funkcyj (2).

Zanim zajmiemy się tym przypadkiem, udowodnimy następującą własność ciągu Sturma:

Lemat. *Żadne dwa kolejne wyrazy ciągu* (2) *nie mogą jednocześnie być równe zeru.*

Dowód. Ze wzorów (1) wynika, że jeżeli dla którejkolwiek z liczb $j=1,2,\dots,k-1$ zachodzą równości $f_{j-1}(\xi)=0$ i $f_j(\xi)=0$, to zachodzi też równość $f_{j+1}(\xi)=0$, skąd wnosimy przez indukcję, że gdyby przy pewnym j z ciągu $1,2,\dots,k-1$ było $f_{j-1}(\xi)=0$ oraz $f_j(\xi)=0$, to musiałoby też być $f_k(\xi)=0$, co nie jest możliwe, gdyż $f_k(x)$ jest stałą różną od zera.

Przy pomocy udowodnionego lematu zbadamy obecnie, jak się zachowuje funkcja $Z(x)$, gdy x przechodzi przez pierwiastek x_0 unkcji $f_j(x)$, gdzie j jest jedną z liczb $0,1,\dots,k-1$.

Niech będzie naprzód $j=0$, t.j. niech x przechodzi przez pierwiastek x_0 wielomianu $f_0(x)=f(x)$. Jak wiemy (Rozdział VIII, § 2, wzór (6), str. 106), mamy tożsamość:

$$f(x)-f(x_0)=(x-x_0)\,Q(x,x_0),$$

zatem, wobec $f(x_0)=0$:

$$f(x)=(x-x_0)\,Q(x,x_0),\tag{3}$$

gdzie

$$Q(x,x)=f'(x)=f_1(x).\tag{4}$$

Ponieważ w myśl lematu nie może być jednocześnie $f(x_0)=0$ i $f_1(x_0)=0$, więc wobec $f(x_0)=0$ jest $f_1(x_0)\neq 0$. Wskutek tego dla x dostatecznie bliskich liczby x_0 wartość $f_1(x)$ ma ten sam znak co wartość $f_1(x_0)$. Ponieważ zaś wobec (4) jest $Q(x_0,x_0)=f_1(x_0)$, więc dla x dostatecznie bliskich x_0 wielomian $Q(x,x_0)$ ma ten sam znak co $f_1(x_0)$, a więc też ten sam co $f_1(x)$. Lecz dla $x<x_0$ mamy $x-x_0<0$, zaś dla $x>x_0$ mamy $x-x_0>0$. Z tożsamości (3) wynika więc, że dla $x<x_0$ znaki liczb $f(x)$ i $Q(x,x_0)$ są różne, zaś dla $x>x_0$ jednakowe. Ostatecznie więc wnosimy, że dla x dostatecznie bliskich x_0, ale mniejszych od x_0, znak liczby $f(x)$ jest różny od znaku liczby $f_1(x)$, zaś dla x dostatecznie bliskich x_0, ale większych od x_0, znaki liczb $f(x)$ i $f_1(x)$ są jednakowe.

Znaczy to, że dla x dostatecznie bliskich x_0 przed przejściem przez x_0 wyrazy $f(x)$ i $f_1(x)$ dawały zmianę znaku, zaś po przejściu przez x_0 już jej nie dawały. A więc w części ciągu (2), utworzonej z pierwszych dwuch jego wyrazów, przy przejściu przez x_0 ubyła jedna zmiana znaku.

Niech teraz j oznacza jedną z liczb ciągu $1,2,\dots,k-1$. Zmiana znaku funkcji $f_j(x)$ może wpłynąć na liczbę zmian znaku tylko w części ciągu Sturma, złożonej z trzech kolejnych funkcyj:

$$f_{j-1}(x),\quad f_j(x),\quad f_{j+1}(x),$$

gdyż zmiana znaku liczby $f_j(x)$ może nastąpić tylko wtedy, gdy x przechodzi przez pierwiastek wielomianu $f_j(x)$, a na mocy lematu żadne dwa kolejne wyrazy ciągu Sturma nie mogą być jednocześnie równe 0.

Niech x_0 oznacza pierwiastek wielomianu $f_j(x)$. Jest więc $f_j(x_0)=0$ i przeto $f_{j-1}(x_0)\neq 0$ oraz $f_{j+1}(x_0)\neq 0$. Zatem $f_{j-1}(x)$ ma dla x dostatecznie bliskich x_0 ten sam znak co $f_{j-1}(x_0)$, zaś $f_{j+1}(x)$ ma ten sam znak co $f_{j+1}(x_0)$. Lecz wobec (1) mamy:

$$f_{j-1}(x)=q_j(x)f_j(x)-f_{j+1}(x),$$

zatem wobec $f_j(x_0)=0$:

$$f_{j-1}(x_0)=-f_{j+1}(x_0),$$

czyli liczby $f_{j-1}(x_0)$ i $f_{j+1}(x_0)$ mają znaki różne. Wnosimy stąd, że dla x dostatecznie bliskich x_0 liczby $f_{j-1}(x)$ i $f_{j+1}(x)$ też mają znaki różne.

Co się tyczy funkcji $f_j(x)$, to może ona przy przejściu x przez x_0 zmieniać znak lub nie. Skoro jednak dla x dostatecznie bliskich x_0 znaki liczb $f_{j-1}(x)$ i $f_{j+1}(x)$ są różne, to dla takich x bez względu na to, jaki jest znak liczby $f_j(x)$ — i nawet jeśli jest ona równa 0 — ciąg trzech liczb:

(5) $$f_{j-1}(x),\quad f_j(x),\quad f_{j+1}(x)$$

daje jedną zmianę znaku. Wyrazy tego ciągu mogą bowiem przyjmować jedną z następujących 6 kombinacyj znaków:

$$++-,\quad +--,\quad +0-,\quad -++,\quad --+,\quad -0+.$$

Przejście wielomianu $f_j(x)$ przez zero nie wywiera zatem żadnego wpływu na liczbę zmian znaku ciągu (2) w części tego ciągu, utworzonej z trzech wyrazów (5). Ponieważ jest to prawdą dla $j=1,2,\dots,k-1$, więc z uwagi na to, że ostatni wyraz $f_k(x)$ ciągu Sturma ma wartość stałą, a więc i znak stały, dochodzimy do wniosku, że zmiana wartości liczby $Z(x)$ zachodzi jedynie wtedy, gdy x przechodzi (rosnąc) przez pierwiastek wielomianu $f(x)$ i że przy każdym takim przejściu ubywa jedna zmiana znaku w ciągu Sturma, czyli że $Z(x)$ zmniejsza się o jedność. Przejście zaś x przez pierwiastek jednego lub więcej wielomianów $f_1(x), f_2(x),\dots,f_{k-1}(x)$ — o ile nie jest połączone z przejściem przez pierwiastek wielomianu $f(x)$ — powodować może jedynie zmianę rozkładu znaków w ciągu Sturma, nie wpływając na wartość liczby $Z(x)$.

Ostatecznie więc, jeżeli $a<b$ oraz $f(a)\neq 0$ i $f(b)\neq 0$ to przy przejściu x od a do b funkcja $Z(x)$ zmniejszy się o tyle jedności, ile jest pierwiastków wielomianu $f(x)$ leżących między a i b. Możemy zatem wypowiedzieć następujące ***twierdzenie Sturma***[1]):

Twierdzenie 1. *Jeżeli wielomian $f(x)$ o współczynnikach rzeczywistych nie posiada pierwiastków wielokrotnych, to liczba pierwiastków wielomianu $f(x)$ leżących między liczbami rzeczywistymi a i $b>a$, które same nie są jego pierwiastkami, jest równa $Z(a)-Z(b)$ czyli ubytkowi liczby zmian znaku przy przejściu od a do b w ciągu Sturma, utworzonym dla wielomianu $f(x)$.*

Tak np. w przykładzie 1 dla wielomianu $f(x)=x^3-x$ mieliśmy $Z(-2)=3$ i $Z(2)=0$, a że -2 i 2 nie są pierwiastkami tego wielomianu, więc na mocy twierdzenia Sturma ma on $Z(-2)-Z(2)=3$ pierwiastki, leżące w przedziale o końcach -2 i 2. Pierwiastkami tymi są liczby -1, 0 i 1. Natomiast, mimo, że $Z(-2)-Z(1)=3-0=3$, nie możemy wnioskować, że ten sam wielomian ma 3 pierwiastki leżące między liczbami $a=-2$ i $b=1$, gdyż mamy tu $f(1)=0$, a więc nie jest zachowane założenie twierdzenia, że liczby a i b same nie są pierwiastkami wielomianu $f(x)$.

Podobnie, w przykładzie 2 wielomian $f(x)=x^5-3x^3+2x$ miał $Z(-2)-Z(2)=5-0=5$ pierwiastków leżących między -2 a 2 (są nimi liczby $-\sqrt{2}$, -1, 0, 1 i $\sqrt{2}$).

---

[1]) Twierdzenie to zakomunikował Akademii Nauk w Paryżu Ch. Sturm w 1829 r. J. A. Serret nazywa je w swej książce *Cours d'Algèbre Supérieure*, Tom I, Wyd. 5-te, Paryż 1885, str. 277, jednym z najwspanialszych odkryć, jakimi wzbogaciła się Analiza matematyczna.

Uwagi. Przy praktycznym stosowaniu twierdzenia Sturma, dogodnie jest korzystać z następujących uwag:

1°. Ponieważ znaki wyrazów ciągu Sturma nie zmienią się przez pomnożenie ich przez liczby dodatnie, więc przy wykonywaniu kolejnych dzieleń dla otrzymania ciągu Sturma możemy, o ile to jest wygodne (np. dla zniesienia mianowników), mnożyć wyrazy tego ciągu przez dowolne liczby dodatnie. Nie wolno natomiast mnożyć przez liczby ujemne, ani opuszczać mianowników ujemnych.

2°. Ponieważ w dowodzie twierdzenia Sturma korzystaliśmy jedynie z tej własności funkcji $f_k(x)$, że jest ona stałego znaku w rozpatrywanym przedziale, ale nie powoływaliśmy się poza tym na tę silniejszą jej własność, że jest ona liczbą stałą różną od 0, więc przy wyznaczaniu ciągu Sturma możemy zatrzymać się na takim wyrazie, o którym wiemy, że nie zmienia znaku w badanym przedziale, choćby nie był stałą.

§ 2. Wnioski z twierdzenia Sturma. Z twierdzenia Sturma wynika natychmiast następujące

Twierdzenie 2. *Jeżeli każdy z pierwiastków rzeczywistych wielomianu o współczynnikach rzeczywistych*

$$f(x)=a_0x^m+a_1x^{m-1}+\ldots+a_m$$

jest większy od a i zarazem mniejszy od b, to liczba wszystkich pierwiastków rzeczywistych tego wielomianu jest równa $Z(a)-Z(b)$.

Otóż takie liczby a i b łatwo znaleźć. Istotnie, oznaczmy przez ϱ dowolną liczbę spełniającą warunki:

$$\varrho\geqslant 1 \quad \text{i} \quad \varrho>\frac{|a_1|+|a_2|+\ldots+|a_m|}{|a_0|}$$

i przyjmijmy $a=-\varrho$, $b=\varrho$. Jeżeli x nie leży między a i b, to musi być $|x|\geqslant\varrho$, zatem $|x|\geqslant 1$ i $|a_0||x|>|a_1|+|a_2|+\ldots+|a_m|$, a przeto:

$$|a_1x^{m-1}+a_2x^{m-2}+\ldots+a_m|\leqslant|x^{m-1}|(|a_1|+|a_2|+\ldots+|a_m|)<$$
$$<|x^{m-1}||x||a_0|=|a_0x^m|,$$

skąd wynika, że x nie może być pierwiastkiem wielomianu $f(x)$ (gdyż wówczas byłoby $a_1x^{m-1}+a_2x^{m-2}+\ldots+a_m=-a_0x^m$, zatem $|a_1x^{m-1}+a_2x^{m-2}+\ldots+a_m|=|a_0x^m|$).

Twierdzenie Sturma daje więc sposób obliczenia liczby wszystkich pierwiastków rzeczywistych wielomianu o współczynnikach rzeczywistych; daje też możność zbadania, czy takie pierwiastki istnieją; równość $Z(-\varrho)=Z(\varrho)$ dowodzi bowiem, że badany wielomian pierwiastków rzeczywistych nie posiada.

Liczbę wszystkich pierwiastków rzeczywistych wielomianu $f(x)$ można jednak otrzymać z ciągu Sturma prościej, a mianowicie w następujący sposób.

Dla x dostatecznie wielkich co do wartości bezwzględnej, każdy z wielomianów ciągu Sturma ma, jak wiemy, taki sam znak jak jego wyraz z najwyższą potęgą x. Wnosimy stąd, że dla obliczenia $Z(-\varrho)$ czy $Z(\varrho)$ nie potrzeba znać liczby ϱ, ani podstawiać jej do ciągu (2), lecz wystarczy obliczyć liczbę zmian znaku w ciągu utworzonym z najwyższych wyrazów wielomianów (2) dla $x<0$ lub dla $x>0$ (więc np. dla $x=-1$ lub $x=1$).

Uwaga ta pozwala zarazem rozstrzygnąć, kiedy wszystkie pierwiastki wielomianu $f(x)$, nie posiadającego pierwiastków wielokrotnych, są rzeczywiste.

Ciąg (2), w którym pierwszy wyraz jest wielomianem stopnia m, a każdy następny (jako reszta z dzielenia) jest stopnia o jedność niższego niż poprzedni, ma co najwyżej $m+1$ wyrazów, a więc może wykazywać co najwyżej m zmian znaku. Jest więc $0\leqslant Z(x)\leqslant m$ przy wszelkim x.

Jeżeli wszystkie m pierwiastków wielomianu $f(x)$ są rzeczywiste, to musi być $Z(-\varrho)-Z(\varrho)=m$, zatem wobec $Z(-\varrho)\leqslant m$ i $Z(\varrho)\geqslant 0$ znajdujemy $Z(-\varrho)=m$ i $Z(\varrho)=0$. Ostatnia równość dowodzi wobec uwagi uczynionej co do obliczania $Z(\varrho)$, że dla $x=1$ najwyższe wyrazy wielomianów (2) mają wszystkie ten sam znak; innymi słowy, że współczynniki przy najwyższych potęgach x wielomianów (2) są wszystkie tego samego znaku. Przy tym, wobec $Z(-\varrho)=m$, ciąg (2) musi mieć co najmniej $m+1$ wyrazów, czyli musi być $k\geqslant m$, a że $k\leqslant m$ na mocy określenia, więc musi tu być $k=m$.

Na odwrót, jeżeli $k=m$ i jeżeli współczynniki przy najwyższych potęgach wielomianów (2) są wszystkie tego samego znaku, to dla ujemnych x dostatecznie wielkich co do wartości bezwzględnej będzie $Z(x)=m$. Skoro bowiem ciąg (2) zawiera $k+1=m+1$ wyrazów, to muszą one być wielomianami odpowiednio stopni:

$$m,\qquad m-1,\qquad m-2,\qquad \ldots,\qquad 1,\qquad 0$$

i najwyższe ich wyrazy, jako mające współczynniki jednakowego znaku, mają dla $x=-1$ znaki odpowiednio takie jak:

$$(-1)^m a_0, \quad (-1)^{m-1} a_0, \quad \ldots, \quad (-1) a_0, \quad a_0;$$

a w ciągu tym jest zawsze m zmian znaku. Udowodniliśmy zatem

Twierdzenie 3. *Na to, żeby wszystkie pierwiastki wielomianu $f(x)$ stopnia m o współczynnikach rzeczywistych, nie mającego pierwiastków wielokrotnych, były rzeczywiste, potrzeba i wystarcza, by odpowiedni ciąg Sturma składał się z $m+1$ wielomianów, w których współczynniki przy najwyższych potęgach x są wszystkie tego samego znaku.*

PRZYKŁAD. Zastosujmy twierdzenie 3 do wyznaczenia warunku koniecznego i dostatecznego na to, aby wszystkie pierwiastki wielomianu

$$f(x) = x^3 + px + q,$$

gdzie liczby p i q są rzeczywiste, były rzeczywiste.

Gdyby wielomian ten miał pierwiastek podwójny, wszystkie jego pierwiastki byłyby rzeczywiste.

Załóżmy więc że nie ma on pierwiastków wielokrotnych.

Mamy tu:

$$f_1(x) = 3x^2 + p \quad \text{i} \quad f_2(x) = (-2px/3) - q.$$

Gdyby było $p=0$, ciąg Sturma dla wielomianu $f(x)$ składałby się tylko z trzech wyrazów i w myśl twierdzenia 3 nie wszystkie pierwiastki jego byłyby rzeczywiste. Załóżmy więc, że $p \neq 0$. Znajdziemy wówczas $f_3(x) = (-27q^2/4p^2) - p$. W myśl twierdzenia 3 warunkiem koniecznym i dostatecznym na to, aby wszystkie pierwiastki wielomianu $f(x)$ były rzeczywiste, będą więc nierówności:

$$-\frac{2p}{3} > 0 \quad \text{i} \quad -\frac{27q^2}{4p^2} - p > 0.$$

Druga z tych nierówności pociąga za sobą pierwszą, wobec czego warunek ten wyraża się przez jedną tylko nierówność drugą (por. Rozdział X, § 3, str. 178):

$$4p^3 + 27q^2 < 0.$$

§ 3. Oddzielanie i przybliżone obliczanie pierwiastków. Twierdzenie Sturma pozwala obliczać pierwiastki rzeczywiste wielomianu o współczynnikach rzeczywistych z dowolną dokładnością. W tym celu należy przede wszystkim dany wielomian zastąpić przez wielomian $f(x)$, mający te same pierwiastki co dany, ale wszystkie jako jednokrotne, co — jak wiemy z twierdzenia 14 Rozdziału VIII (§ 10, str. 122) — jest zawsze możliwe.

Następnie należy wyznaczyć (np. w sposób, jaki wskażemy na str. 274) przedział (a,b), wewnątrz którego znajdują się wszystkie pierwiastki rzeczywiste wielomianu $f(x)$. Jeżeli teraz chcemy obliczyć pierwiastki wielomianu $f(x)$ z dokładnością do danej liczby dodatniej ε, to wystarczy podzielić przedział (a,b) na przedziały o długości mniejszej niż ε, np. na n równych przedziałów, gdzie $n>(b-a)/\varepsilon$, zbadać (przez podstawienie), czy końce tych przedziałów nie są pierwiastkami wielomianu $f(x)$ — co ewentualnie dałoby dokładne wartości pewnej liczby pierwiastków — i, stosując twierdzenie Sturma, wyznaczyć, ile pierwiastków wielomianu $f(x)$ znajduje się wewnątrz każdego z tych przedziałów. Dla każdego z pierwiastków, znajdujących się wewnątrz któregokolwiek z tych przedziałów, końce tego przedziału dadzą wartości pierwiastka z błędem mniejszym od ε.

Mamy tu więc, teoretycznie przynajmniej, pewną metodę obliczania z dowolną dokładnością pierwiastków rzeczywistych każdego równania o danych współczynnikach rzeczywistych.

W praktyce sposób ten może jednak wymagać uciążliwych rachunków; praktyczne rozwiązywanie równań algebraicznych stanowi przedmiot odrębnych badań, którymi się tu zajmować nie będziemy [1]). Ograniczymy się tylko do uwag następujących.

Przede wszystkim możemy poprzestać na znajdowaniu pierwiastków dodatnich wielomianu $f(x)$, gdyż pierwiastki ujemne tego wielomianu są oczywiście pierwiastkami dodatnimi wielomianu $f(-x)$ i na odwrót, zaś 0 jest wtedy i tylko wtedy pierwiastkiem wielomianu $f(x)$, gdy jego wyrazem wolnym jest 0. Otóż przedział, wewnątrz którego mieszczą się wszystkie pierwiastki dodatnie wielomianu $f(x)$, znaleźć można, jak następuje.

Niech

(6) $$f(x) = a_0 x^m + a_1 x^{m-1} + \ldots + a_m.$$

---

[1]) Por. cytowaną na str. 103 książkę C. Runge'go, *Praxis der Gleichungen*.

Możemy oczywiście założyć, że $a_0>0$, gdyż w przeciwnym razie możnaby zastąpić wielomian $f(x)$ przez wielomian $-f(x)$. Gdyby wszystkie współczynniki $a_0, a_1, \ldots, a_m$ były dodatnie, równanie $f(x)=0$ nie posiadałoby oczywiście pierwiastków dodatnich. Załóżmy więc, że tak nie jest. Niech a_k (gdzie k jest jedną z liczb ciągu $1, 2, \ldots, m$) oznacza pierwszy współczynnik ujemny wielomianu $f(x)$, zaś ω bezwzględną wartość największego (co do bezwzględnej wartości) lub jednego z jednakowych największych (co do bezwzględnej wartości) ujemnych współczynników wielomianu (6). Mamy więc:

$$a_0>0,\quad a_1\geqslant 0,\quad \ldots,\quad a_{k-1}\geqslant 0,\quad a_k\geqslant -\omega,\quad a_{k+1}\geqslant -\omega,\quad \ldots,\quad a_m\geqslant -\omega,$$

skąd dla $x>0$:

$$f(x)\geqslant a_0x^m-\omega x^{m-k}-\omega x^{m-k-1}-\ldots-\omega x-\omega=$$

$$=a_0x^m-\omega\frac{x^{m-k+1}-1}{x-1}=x^{m-k+1}\frac{a_0x^{k-1}(x-1)-\omega}{x-1}+\frac{\omega}{x-1},$$

zatem dla $x>1$ wobec $\omega>0$:

$$f(x)\geqslant x^{m-k+1}\frac{a_0x^{k-1}(x-1)-\omega}{x-1}\geqslant x^{m-k+1}\frac{a_0(x-1)^k-\omega}{x-1}. \tag{7}$$

Dla $x\geqslant 1+\sqrt[k]{\dfrac{\omega}{a_0}}$ mamy $a_0(x-1)^k\geqslant\omega$, zatem wobec (7) $f(x)>0$.

A więc:

Każdy pierwiastek dodatni wielomianu $f(x)$ jest mniejszy od

$$H=1+\sqrt[k]{\frac{\omega}{a_0}}.$$

Tak samo znaleźlibyśmy dolną granicę dodatnią h pierwiastków dodatnich wielomianu $f(x)$, zważywszy, że jest ona równa górnej granicy pierwiastków dodatnich wielomianu $x^mf(1/x)$.

Mając przedział (a,b), w którym zawarte są wszystkie pierwiastki rzeczywiste danego wielomianu (6), dzielimy go na mniejsze przedziały, np. za pomocą przepoławiania, i badamy w myśl twierdzenia Sturma ilość pierwiastków wielomianu $f(x)$ w tych mniejszych przedziałach, odrzucając dalej te, w których pierwiastków niema i t. d.

Powtarzając to postępowanie dostateczną liczbę razy, dojdziemy oczywiście bądź do pierwiastków wielomianu, bądź do przedziałów, wewnątrz których mieści się po jednym tylko jego pierwiastku, czyli dokonamy t. zw. *oddzielenia pierwiastków.*

Nastąpi to w każdym razie wtedy, gdy długość rozważanych przedziałów będzie mniejsza od bezwzględnej wartości różnicy jakichkolwiek dwóch różnych pierwiastków wielomianu. Liczbę taką można wyznaczyć za pomocą współczynników wielomianu.

Istotnie, niech $x_1, x_2, \ldots, x_m$ będą jego pierwiastkami. Współczynniki równania

$$F(t)=\prod_{k=2}^{m}\prod_{l=1}^{k-1}[t-(x_k-x_l)^2]=0$$

są oczywiście funkcjami symetrycznymi pierwiastków wielomianu $f(x)$, a więc w myśl twierdzenia Cauchy'ego (Rozdział IX, § 3, str. 159, tw. 2) dadzą się wyrazić jako wielomiany względem jego współczynników. Jeżeli teraz oznaczymy przez h dolną granicę dodatnią pierwiastków wielomianu $F(t)$, które są wszystkie dodatnie (znalezioną jak podano wyżej), to różnica bezwzględna dwóch różnych pierwiastków wielomianu $f(x)$ będzie oczywiście większa od $\sqrt{h}$. Jeżeli więc przedział, w którym mieszczą się wszystkie pierwiastki rzeczywiste wielomianu $f(x)$, podzielimy na przedziały o długości co najwyżej $\sqrt{h}$, to dokonamy oddzielenia pierwiastków wielomianu $f(x)$.

Gdy pierwiastki wielomianu $f(x)$ są oddzielone, to przy dalszym rozdrabnianiu przedziałów nie trzeba już badać ciągu Sturma, lecz wystarczy uwzględniać tylko znak funkcji dla punktów podziału danego przedziału.

Istotnie, niech w przedziale (p,q) leży tylko jeden pierwiastek x_0 wielomianu $f(x)$. Wielomian $f(x)$ jest podzielny przez $x-x_0$, przy czym iloraz $Q(x)$ z tego dzielenia nie jest zerem dla $x=x_0$, gdyż—jak zakładamy—wielomian $f(x)$ nie ma pierwiastków wielokrotnych. Zatem $f(x)=(x-x_0)Q(x)$, gdzie $Q(x_0)\neq 0$, i przeto przy przejściu przez x_0 wielomian $f(x)$ zmienia znak (gdyż $x-x_0<0$ dla $x<x_0$, zaś $x-x_0>0$ dla $x>x_0$). Znaki liczb $f(p)$ i $f(q)$ muszą więc być różne, gdyż w przeciwnym razie wielomian $f(x)$ zmieniałby wewnątrz przedziału (p,q) znak więcej niż jeden raz i przeto, jako funkcja ciągła, przechodziłby wewnątrz przedziału (p,q) przez zero więcej niż jeden raz, wbrew założeniu, że w przedziale tym jest tylko jeden pierwiastek.

Niech więc będzie np. $f(p)<0$ i $f(q)>0$. Niech teraz r oznacza jakąkolwiek liczbę, leżącą wewnątrz przedziału (p,q). Jeżeli przypadkowo $f(r)=0$, to znaleźliśmy pierwiastek wielomianu $f(x)$, jeżeli zaś $f(r)>0$, to pierwiastek wielomianu $f(x)$, leżący między p i q, będzie leżał między p i r; jeżeli wreszcie $f(r)<0$, to będzie on leżał między r i q.

§ 4. *Regula falsi* i metoda Newtona. Są różne sposoby wybierania nowych punktów podziału przy rozdrabnianiu przedziałów, zawierających pierwiastek. Omówimy tu szkicowo dwa najważniejsze z nich, bez bliższego roztrząsania warunków, przy jakich są one skuteczne, t.j. przy jakich pozwalają szybko obliczyć pierwiastek z dużą dokładnością.

1^0 *Regula falsi.* Niech pierwiastek x_0 wielomianu (6) zawiera się wewnątrz małego przedziału (p,q). Jest więc $p<x_0<q$, przy czym liczby $x_0-p=\varepsilon_1$ i $q-x_0=\varepsilon_2$ są małe. W myśl wzoru Taylora (Rozdział VIII, § 8, wzór (28), str. 119) mamy:

$$f(p)=f(x_0-\varepsilon_1)=f(x_0)-\frac{\varepsilon_1}{1!}f'(x_0)+\frac{\varepsilon_1^2}{2!}f''(x_0)-\dots,$$

$$f(q)=f(x_0+\varepsilon_2)=f(x_0)+\frac{\varepsilon_2}{1!}f'(x_0)+\frac{\varepsilon_2^2}{2!}f''(x_0)+\dots$$

Mamy $f(x_0)=0$, zaś z uwagi na to, że liczby ε_1 i ε_2 są małe, możemy pominąć drugie i wyższe ich potęgi, przez co otrzymamy przybliżone równości:

$$f(p)=-\varepsilon_1 f'(x_0), \qquad f(q)=\varepsilon_2 f'(x_0).$$

Wobec $f(x_0)=0$, jest $f'(x_0)\neq 0$, gdyż wielomian (6) nie posiada — jak założyliśmy — pierwiastków wielokrotnych. Mamy stąd proporcję przybliżoną:

$$\frac{\varepsilon_1}{\varepsilon_2}=-\frac{f(p)}{f(q)},$$

dającą wobec $\varepsilon_1+\varepsilon_2=q-p$ równość przybliżoną:

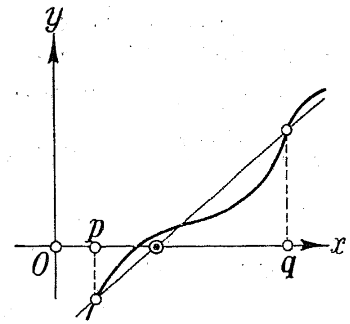
$$x_0=p+\varepsilon_1=\frac{pf(q)-qf(p)}{f(q)-f(p)}.$$

Za przybliżoną wartość pierwiastka, albo też za punkt podziału przedziału (p,q), można więc wziąć liczbę

$$\frac{pf(q)-qf(p)}{f(q)-f(p)}.$$

Względem nowego, mniejszego przedziału można następnie zastosować podobne postępowanie.

Znaczenie geometryczne opisanej metody jest następujące: zamiast odciętej punktu przecięcia się krzywej $y=f(x)$ z osią x-ów, bierzemy za przybliżoną wartość pierwiastka odciętą punktu przecięcia się z osią x-ów cięciwy tej krzywej, łączącej jej punkty o odciętych p i q (Rys. 5).



Rys. 5.

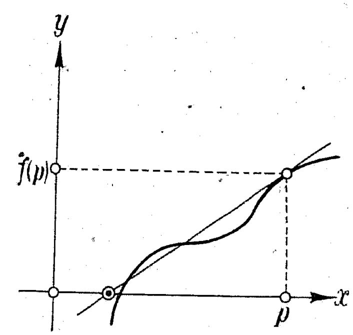
2^0 *Metoda Newtona.* Jeżeli p jest przybliżoną wartością pierwiastka x_0 wielomianu (6), to mamy $x_0=p+\varepsilon$, gdzie ε jest liczbą małą. Pomijając jej drugą i wyższe potęgi w rozwinięciu Taylora

$$f(x_0)=f(p+\varepsilon)=f(p)+\frac{\varepsilon}{1!}f'(p)+\frac{\varepsilon^2}{2!}f''(p)+\dots,$$

otrzymamy wobec $f(x_0)=0$ równość przybliżoną $0=f(p)+\varepsilon f'(p)$, skąd $\varepsilon=-\dfrac{f(p)}{f'(p)}$ (jeżeli $f'(p)\neq 0$). Przybliżoną wartością pierwiastka $x_0=p+\varepsilon$ będzie więc:

$$p-\frac{f(p)}{f'(p)}.$$

Z tą nową wartością przybliżoną możemy postąpić dalej podobnie jak z wartością p i t. d. Oczywiście, jeżeli $f'(p)=0$ lub jeżeli stosunek $f(p):f'(p)$ jest duży, to metoda ta nie daje się stosować.



Rys. 6.

Podobnie jak dla metody *regula falsi* możemy metodę Newtona zinterpretować geometrycznie (Rys. 6): za przybliżoną wartość pierwiastka przyjmujemy w niej odciętą punktu przecięcia osi x i stycznej do krzywej $y=f(x)$ w punkcie $(p, f(p))$.

Dla otrzymania lepszego przybliżenia możnaby też uwzględniać pierwsze trzy składniki rozwinięcia $f(x_0)$ i wyznaczać ε z równania 2-go stopnia

$$f(p)+\varepsilon f'(p)+\frac{\varepsilon^2}{2}f''(p)=0.$$

Zastosujmy metodę **Newtona** do obliczania pierwiastka dodatniego równania $x^2=D$, gdzie $D>1$. Za przybliżoną wartość pierwiastka weźmy $p_0=1$. Następnym przybliżeniem będzie więc $p_1=p_0-\frac{p_0^2-D}{2p_0}=\frac{p_0^2+D}{2p_0}$, dalszym $p_2=\frac{p_1^2+D}{2p_1}$ i t. d.; ogólnie $p_{n+1}=\frac{p_n^2+D}{2p_n}$ dla $n=1,2,\ldots$ Przyjmijmy jeszcze:

$$q_n=\frac{D}{p_n} \quad \text{dla} \quad n=0,1,2,\ldots$$

Będziemy mieli:

$$p_{n+1}=\frac{p_n+q_n}{2}, \quad q_{n+1}=\frac{2p_nq_n}{p_n+q_n} \quad \text{dla} \quad n=0,1,2,\ldots$$

Metoda **Newtona** doprowadza więc do znanej z kursu Analizy metody obliczania pierwiastków kwadratowych za pomocą średnich arytmetyczno-harmonicznych, dającej — jak wiadomo — ciągi bardzo szybko zbieżne.

PRZYKŁAD. Obliczmy pierwiastki dodatnie wielomianu

$$f(x)=x^3+7x-7.$$

Mamy tu $k=2$, $\omega=7$, $1+\sqrt[k]{\frac{\omega}{a_0}}=1+\sqrt{7}<4$. Pierwiastki dodatnie wielomianu $f(x)$ są więc mniejsze od 4. Dla obliczenia, ile ich jest, tworzymy ciąg **Sturma**. Mamy tu $f_1(x)=3x^2+7$, skąd dalej znajdujemy $f_2(x)=-\frac{14}{3}x+7$. Wielomian ten możemy zastąpić jego iloczynem przez liczbę dodatnią 3/7, t. j. przez wielomian $-2x+3$. Dzieląc $f_1(x)$ przez $-2x+3$, otrzymujemy resztę 55/4. Zatem $f_3(x)$ jest stałą ujemną.

Dla $x=0$ ciąg **Sturma** ma znaki $-$, $+$, $+$, $-$, t. j. dwie zmiany znaku, zaś dla $x=4$ znaki $+$, $+$, $-$, $-$, t. j. jedną zmianę znaku. Wnosimy stąd, że między 0 a 4 mieści się jeden pierwiastek wielomianu $f(x)$.

Ponieważ $f(2)>0$, więc wobec $f(0)<0$ pierwiastek ten mieści się wewnątrz przedziału $(0,2)$. Ponieważ dalej $f(1)>0$, więc pierwiastek ten leży wewnątrz przedziału $(0,1)$.

Biorąc $p=1$ i stosując metodę **Newtona**, otrzymamy jako przybliżoną wartość pierwiastka liczbę:

$$1-\frac{f(1)}{f'(1)}=1-\frac{1}{10}=0,9.$$

Ponieważ $f(0,9)>0$, zaś $f(0,89)<0$, więc pierwiastek ten leży między 0,89 a 0,9. Liczba 0,9 daje więc jego wartość z błędem mniejszym od 0,01.

§ 5. Obliczanie pierwiastków zespolonych wielomianu o dowolnych współczynnikach zespolonych. Niech $f(z)$ będzie wielomianem o dowolnych współczynnikach zespolonych:

$$f(z)=a_0z^m+a_1z^{m-1}+\ldots+a_m,$$

gdzie $a_k=b_k+c_ki$ dla $k=0,1,2,\ldots,m$, przy czym liczby b_k i c_k są rzeczywiste.

Dla liczby zespolonej $z=x+yi$ mamy:

$$f(z)=f(x+yi)=\sum_{k=0}^{m}(b_k+c_ki)(x+yi)^{m-k}=P(x,y)+iQ(x,y),$$

gdzie $P(x,y)$ i $Q(x,y)$ są wielomianami (dwóch zmiennych x i y) o współczynnikach rzeczywistych. Na to, żeby było $f(z)=0$, potrzeba i wystarcza, aby było jednocześnie:

$$P(x,y)=0 \quad \text{i} \quad Q(x,y)=0.$$

Obliczanie pierwiastków zespolonych wielomianu $f(z)$ sprowadza się w ten sposób do obliczania pierwiastków rzeczywistych, wspólnych dla dwu wielomianów dwu zmiennych o współczynnikach rzeczywistych, co znowu — jak wiemy z teorii eliminacji (por. Rozdział XV, § 3, str. 257) — sprowadza się do wyznaczania pierwiastków wielomianów jednej zmiennej o współczynnikach rzeczywistych. W ten sposób ostatecznie obliczanie pierwiastków zespolonych dowolnego wielomianu o współczynnikach zespolonych sprowadza się do obliczania pierwiastków rzeczywistych wielomianu o współczynnikach rzeczywistych.

Zauważymy tu jeszcze, że — jak dowiódł **Carmichael**[1]) — każdy pierwiastek ζ wielomianu (6) przy $a_0=1$ spełnia nierówność:

$$|\zeta|\leqslant|a_1|+\sqrt{|a_2|}+\sqrt[3]{|a_3|}+\ldots+\sqrt[m]{|a_m|}.$$

Dowód tego twierdzenia podał też **Walsh**[2]) oraz profesor **Rudnicki**[3]).

[1]) R. D. Carmichael, Bulletin of the American Mathematical Society, Tom 24 (1917-18), str. 293.

[2]) J. L. Walsh, Annals of Mathematics (2), Tom 25 (1924), str. 285-286.

[3]) J. Rudnicki, Rocznik Polskiego Towarzystwa Matematycznego, Tom 11 (1932), str. 14-18.