STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

REKENAFDELING

MR 105

Difference schemes with complex time steps

by

P.J. van der Houwen

RA

May 1969

Contents

Introduction

In this paper we investigate difference schemes of the type

$$u_{k+1} = (1 + \tau D_0)u_k, \; k = 0, 1, 2, \ldots,$$

where D_0 is a matrix operator not depending on k with purely imaginary eigenvalues and τ is the time step. As an example of such a difference scheme we mention the scheme studied by Lauwerier and Damsté (see reference [3]) for solving the North Sea Problem (without friction) numerically. It was proved that the scheme is unstable and that one had to introduce a relatively large coefficient of friction in the model in order to get a stable scheme with an acceptable step. For realistic values of the coefficient of friction, however, it turned out that the stability condition prescribes a time step which is too small for actual computations. In [2] it was shown that one may obtain stability for schemes of the type given above by introducing non-uniform complex time steps. This gave rise to the following problem:

> Given the class of polynomials $Q_n(z)$ of degree n in z which are of the type
>
> $$Q_n(z) = (1 - \beta_2 z + \beta_4 z^2 - \ldots)^2 + z(1 - \beta_3 z + \beta_5 z^2 - \ldots)^2,$$
>
> where $\beta_2, \ldots, \beta_n$ are real parameters.
> One is asked to determine that polynomial $Q_n(z)$ in this class which is less than or equal to 1 on an interval $0 \leq z \leq b$ where b is as large as possible.

In [2] this problem was solved for n = 2, 3, and 4.

Here, we solve this problem for odd values of n. For even values of n the problem appeared to be far more difficult. Nevertheless, we did succeed in constructing a polynomial with $b(n - 1) < b(n) < b(n + 1)$. Further, we investigate the effect of introducing a linear friction term into the difference scheme.

1. The method of non-uniform complex time steps

Consider the difference scheme

(1.1) $$u_{k+1} = (1 + \tau D)u_k, \; k = 0, 1, 2, \ldots,$$

where the functions u_k are grid functions defined on a discrete set of points, τ is a positive parameter not depending on k and D is a linear difference operator which also does not depend on k. When u_0 is given one can construct the functions $u_1, u_2, \ldots$ with the recurrence relation (1.1). An important class of linear initial boundary value problems may be approximated by difference schemes of type (1.1) (see reference [2], p. 27).

The parameter τ, which we shall call the time step of the difference scheme, is determined by imposing some stability condition upon the scheme. For instance, one may require that all eigenvalues of the operator $1 + \tau D$ are within or on the unit circle (compare [2], p. 30). Such a condition leads to the following upper bound for τ:

(1.2) $$\tau \leq - 2 \underset{\delta_j}{\mathrm{Min}} \frac{\mathrm{Re}\; \delta_j}{|\delta_j|^2},$$

where δ_j are the eigenvalues of D.
Clearly, there is no stability when D has one or more purely imaginary eigenvalues.

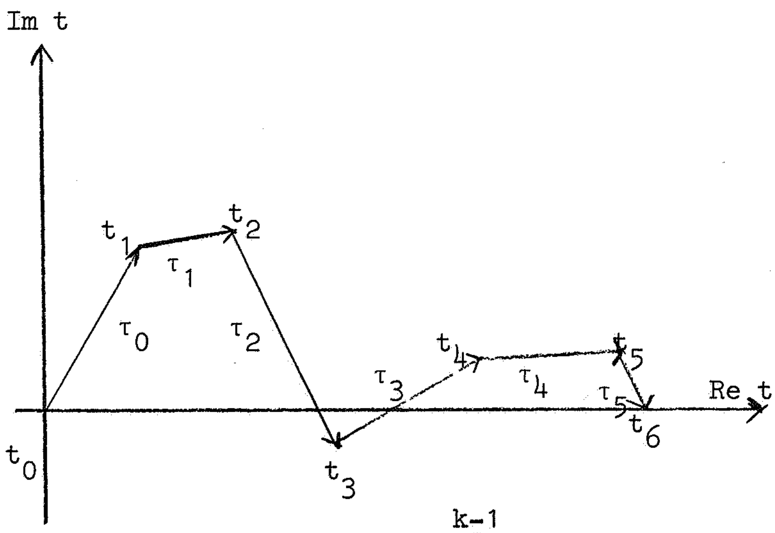
In this paper we shall mainly be interested in schemes of type (1.1) where $D = D_0$ has imaginary eigenvalues, i.e.

(1.3) $$\delta_j = iy_j \;, \quad -\sigma(D_0) \leq y_j \leq \sigma(D_0),$$

$\sigma(D_0)$ being the spectral radius of D_0. The problem is to modify (1.1) in such a way that the scheme becomes stable.
In [2], p. 45 it was proposed to drop the condition that τ is a constant positive time step, but to allow that τ has non-uniform complex values, i.e.

(1.4) $$u_{k+1} = (1 + \tau_k D_0)u_k, \; k = 0, 1, 2, \ldots .$$

This means that we have extended the analytical solution $\tilde{U}(t)$ over complex values of the time t and that this solution is calculated not along the positive time axis but along some path in the complex time plane (see figure 1.1). Of course, this path has to be chosen in such a way that there are enough points t_k on the positive t-axis in order to give information about the behaviour of the solution for real values of t. We shall return to this subject in section 6.



Complex t-plane, $t_k = \sum_{j=0}^{k-1} \tau_j$

fig. 1.1

Let us fix some value of k, say k = n, and let us consider the expression

$$(1.5)\qquad u_n = (1 + \tau_{n-1}D_0)u_{n-1} = \prod_{j=0}^{n-1} (1 + \tau_j D_0)u_0 = P_n(D_0)u_0.$$

$P_n(D_0)$ is a polynomial of degree n in D_0. It can be written as

$$(1.6)\qquad P_n(D_0) = 1 + \tau D_0 + \beta_2 \tau^2 D_0^2 + \ldots + \beta_n \tau^n D_0^n,$$

where

$$\tau = \tau_0 + \tau_1 + \ldots + \tau_{n-1}, \ \ldots, \ \beta_n = \frac{\tau_0 \tau_1 \cdots \tau_{n-1}}{\tau^n} .$$

The coefficients $\beta_2, \ldots, \beta_n$ are dimensionless constants. We shall require that the total time step τ and the coefficients $\beta_2, \ldots, \beta_n$ are real numbers.

We now apply the stability criterion to (1.5):

$$|P_n(\delta_j)| \leq 1 \quad \text{for all} \quad \delta_j. \tag{1.7}$$

This condition can be formulated more explicitly. Introducing the variable $z = \tau^2 y^2$ we obtain for $|P_n(\delta_j)|$ the expression

$$|P_n(\delta_j)| = \sqrt{Q_n(z_j)}\ , \quad j = 1, 2, \ldots \tag{1.8}$$

$$Q_n(z) = (1 - \beta_2 z + \beta_4 z^2 - \ldots)^2 + z(1 - \beta_3 z + \beta_5 z^2 - \ldots)^2. \tag{1.9}$$

Further, let us define the number b as the maximal number such that $Q_n(z)$ is less than or equal to 1 for $0 \leq z \leq b$. Then condition (1.7) reduces to an upper bound for the time step τ, namely

$$\tau \leq \frac{\sqrt{b}}{\sigma(D_0)}\ . \tag{1.7'}$$

The value of b is a function of n and the vector $\vec{\beta} = (\beta_2, \ldots, \beta_n)$, thus $b = b(n,\vec{\beta})$. In actual computation we are interested in the maximal value of b for a given value of n. Thus we are looking for the value of

$$b(n) = \underset{\vec{\beta}}{\mathrm{Max}}\ b(n,\vec{\beta}) \tag{1.10}$$

and the vector $\vec{\beta}$ for which this maximum is reached.

Finally, we remark that in order to compare the effectiveness of the difference schemes for different values of n we must introduce the effective time step

$$\tau_{eff} = \frac{\tau}{n} \leq \frac{\sqrt{b(n)}}{n\sigma(D_0)}\ . \tag{1.11}$$

2. Some direct results

In [2] p. 46 ff. the cases n = 2, 3 and 4 were treated by a straightforward analysis. In table 2.1 the results of this analysis are collected. In addition, we have listed the trivial cases n = 0, 1 and the limiting case $n = \infty$. The latter case needs some explanation. For $n = \infty$ the polynomials between brackets in the formula for $Q_n(z)$ are infinite Taylor expansions of arbitrary functions c(x) and s(x) satisfying the condition c(0) = s(0) = 1, i.e.

$$Q_\infty(z) = c^2(z) + zs^2(z), \; c(0) = s(0) = 1.$$

It is readily seen that $c(z) = \cos\sqrt{z}$ and $s(z) = \sin\sqrt{z}/\sqrt{z}$ make $Q_\infty(z)$ identical to 1 while the condition c(0) = s(0) = 1 is satisfied. Hence $b(\infty) = \infty$ and the coefficients $\beta_2, \ldots, \beta_n$ follow from the Taylor expansions in $\sqrt{z}$ of $\cos\sqrt{z}$ and $\sin\sqrt{z}$, i.e. $\beta_j = 1/j!$.

Table 2.1

n	b(n)	β_2	β_3	β_4	$Q_n(z)$	τ_{eff} $\sigma(D)$
0	∞	0	0	0	1	
1	0	0	0	0	$1 + z$	0
2	1	1	0	0	$1 - z(1 - z)$	.5
3	4	1/2	1/4	0	$1 - z^2(4 - z)/16$	.67
4	8	1/2	1/6	1/24	$1 - z^3(8 - z)/576$	.71
...	...	...	...	...	...	...
∞	∞	$\beta_j = 1/j!$			1	?

3. An upperbound for b(n)

For $m \geq 5$ the direct methods applied in [2] are too complicated to find b(n). In this section a more general approach is given which immediately yields an upperbound for b(n).

First an explicit expression for b(n) is derived. Consider those polynomials $Q_n(z)$ for which $b(n,\vec{\beta}) > 0$ and introduce the new variable

(3.1) $$x = \frac{z}{b(n,\vec{\beta})} .$$

Substituting $z = xb(n,\vec{\beta})$ into $Q_n(z)$ we obtain

(3.2) $$Q_n(z) = C_p^2(x) + b(n,\vec{\beta})\ x\ S_q^2(x),$$

where

$$C_p(x) = 1 - \alpha_2 x + \alpha_4 x^2 - \dots + (-)^p \alpha_{2p} x^p,$$

$$S_q(x) = 1 - \alpha_3 x + \alpha_5 x^2 - \dots + (-)^q \alpha_{2q+1} x^q,$$

$$p = q = \frac{n-1}{2} \text{ if n is odd,}$$

$$p = \frac{n}{2},\ q = p - 1 \text{ if n is even,}$$

and

$$\alpha_{2j} = \beta_{2j} b^j(n,\vec{\beta}),\ \alpha_{2j+1} = \beta_{2j+1} b^j(n,\vec{\beta}),\ j = 1, 2, \dots .$$

Since the transformation $\vec{\beta} \to \vec{\alpha}$ is a 1-1 mapping of the $\vec{\beta}$ space on the $\vec{\alpha}$ space we may consider $b(n,\vec{\beta})$ as a function of n and $\vec{\alpha}$. When we have this function in mind we shall write $b(n,\vec{\alpha})$.
From (3.2) and the condition $Q_n(z) \leq 1$ for $0 \leq z \leq b(n,\vec{\beta})$ or $0 \leq x \leq 1$ we find

(3.3) $$b(n,\vec{\alpha}) \leq \frac{1 - C_p^2(x)}{x\ S_q^2(x)} \equiv B(n,x,\vec{\alpha}),\ 0 \leq x \leq 1.$$

Hence

(3.4) $$b(n,\vec{\alpha}) = \min_{0\leq x\leq 1} B(n,x,\vec{\alpha})$$

and

(3.5) $$b(n) = \max_{\vec{\alpha}} \min_{0\leq x\leq 1} B(n,x,\vec{\alpha}).$$

The problem of finding b(n) is reduced to what we may interpret as a minimax problem for a class of rational functions $B(n,x,\vec{\alpha})$.

We now construct an upperbound for b(n) by majorizing the value of the function $B(n,x,\vec{\alpha})$ in x = 0. In the following section we shall show that for odd values of n there exists a function $B(n,x,\vec{\alpha})$ with b(n) equal to this upperbound. This proves the function to be optimal. From the definition of $B(n,x,\vec{\alpha})$ it follows that

(3.6) $$B(n,0,\vec{\alpha}) = 2\alpha_2.$$

Combining this with (3.3) and (3.5) we see that

(3.7) $$b(n) \leq 2 \operatorname{Maximum}_{\alpha_2, C_p^2(x)\leq 1} \alpha_2.$$

The condition $C_p^2(x) \leq 1$ is equivalent to the assumption that all $b(n,\vec{\beta})$ are non-negative as was required above.
Now $-\alpha_2$ is the derivative of $C_p(x)$ in x = 0, thus we are led to the problem to construct a polynomial $C_p(x)$ of degree p in x with $C_p(0) = 1$ which has a maximal slope in x = 0. In [2] p. 38 it was proved that the transformed Chebyshev polynomial

(3.8) $$T_p(1 - 2x) = \cos(p \arccos (1 - 2x))$$

possesses this property. Further, it was shown that

(3.9) $$\left[\frac{d}{dx} T_p(1 - 2x)\right]_{x=0} = - 2p^2.$$

From (3.7) and (3.9) we obtain the result

$$b(n) \leq 4p^2 = \begin{cases} (n-1)^2 & \text{if } n \text{ is odd} \\ n^2 & \text{if } n \text{ is even} \end{cases} . \tag{3.10}$$

By substituting this into (1.7') and using (1.11) it is seen that

$$\tau_{eff} \leq \frac{1}{\sigma(D)} . \tag{3.11}$$

Hence the fourth degree polynomial operator $P_4(D)$ has an effective time step which is already about 70% of the maximal expectable effective time step. Nevertheless it pays when we should be able to construct a polynomial operator with $\tau_{eff} \sim 1/\sigma(D)$.

4. The optimal polynomial for odd values of n

The considerations in the preceding section suggest the function

(4.1) $$C_p(x) = T_p(1 - 2x),\ p = \frac{n-1}{2}\ .$$

From the properties of the Chebyshev polynomials it follows that the function $1 - C_p^2(x)$ has p+1 zeroes in the interval $0 \leq x \leq 1$ (see figure 4.1).

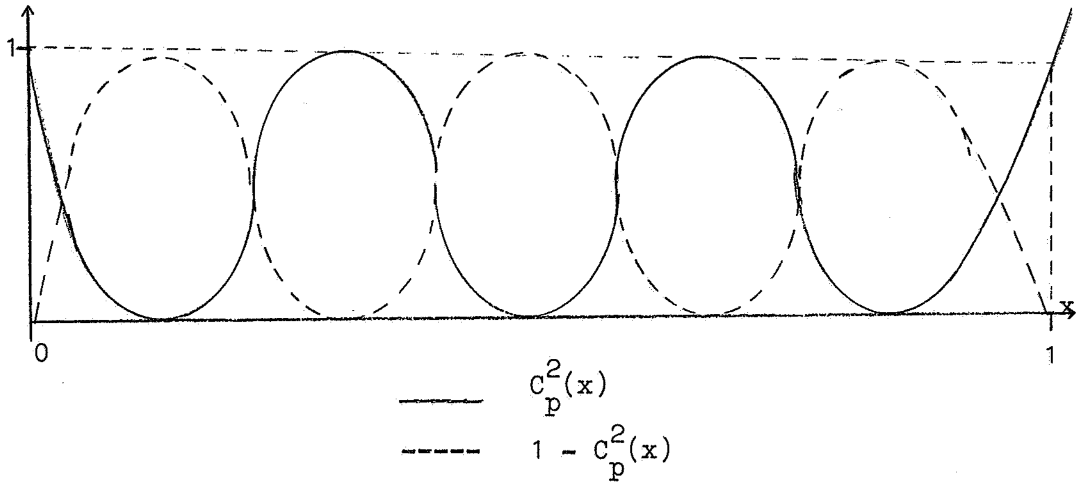


fig. 4.1

In order to obtain a non-zero value for $b(n,\vec{\alpha})$ we have to select the polynomial $S_q(x) = S_p(x)$ in such a way that $xS_p^2(x)$ has the same zeroes as $1 - C_p^2(x)$, i.e.

$$S_p(x) = c(1 - x)\ C_p'(x) \equiv c(1 - x)\ \frac{d}{dx}\ T_p(1 - 2x),$$

where c is a constant determined by the condition $C_p(0) = 1$. From (3.9) it follows that $c = -1/2p^2$, so that

(4.2) $$S_p(x) = -\frac{1-x}{2p^2}\frac{d}{dx}\ T_p(1 - 2x) = \frac{1-x}{p}\ U_{p-1}(1 - 2x),$$

where U_{p-1} is the Chebyshev polynomial of the second kind (see [1], p. 796). It may be remarked that the choice (4.2) is only possible for odd values of n, i.e. if p = q. For even values of n we have q = p - 1 so that $S_q(x)$ has not enough zeroes to neutralize the zeroes of $1 - C_p^2(x)$.

From the definition of U_{p-1} it is easily verified that

$$U_{p-1}(1 - 2x) = \frac{\sin(p \arccos(1 - 2x))}{2\sqrt{x(1 - x)}} . \tag{4.3}$$

By substituting (3.8) and (4.3) into (4.1) and (4.2) we obtain for $B(n,x,\vec{\alpha})$ the expression

$$B(n,x,\vec{\alpha}) = \frac{4p^2}{1 - x} = \frac{(n - 1)^2}{1 - x} . \tag{4.4}$$

Hence

$$b(n,\vec{\alpha}) = (n - 1)^2,$$

and by virtue of (3.10)

$$b(n) = (n - 1)^2. \tag{4.5}$$

The effective time step becomes

$$\tau_{eff} = \frac{1 - \frac{1}{n}}{\sigma(D)} , \tag{4.6}$$

which can be made as close to its asymptotic value as one wishes. Further, the optimal polynomial $Q_n(z)$ assumes the form

$$Q_n(z) = 1 - 4 \frac{z^2(b(n) - z)}{b^3(n)} U^2_{\frac{n-3}{2}}(1 - 2 \frac{z}{b(n)}). \tag{4.7}$$

Finally, the results for the cases n = 5, 7, 9 are listed in table 4.1.

Table 4.1

n	b(n)	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	$\tau_{eff}\ \sigma(D)$
5	16	8/b	3/b	$8/b^2$	$2/b^2$	0	0	0	0	.800
7	36	18/b	19/3b	$48/b^2$	$32/3b^2$	$32/b^3$	$16/3b^3$	0	0	.857
9	64	32/b	11/b	$160/b^2$	$34/b^2$	$256/b^3$	$40/b^3$	$128/b^4$	$16/b^4$	.889

5. Some remarks about the polynomial for even values of n

Although from a practical point of view the solution of the problem for even values of n is not important, it is for reasons of completeness that we have investigated the even case.

First, we observe that for even values of n

$$(5.1) \qquad (n - 2)^2 \le b(n) \le n^2.$$

This is readily seen from (3.10) and (4.5).
Further, it is clear that we cannot choose $C_p(x)$ according to (4.1), because the numerator $xS_q^2(x) = xS_{p-1}^2(x)$ in (3.3) has not enough zeroes to neutralize the zeroes of $1 - C_p^2(x)$, the denumerator.
Let us start with the expression $B(n-1,x,\vec{\alpha})$, where n is a given even number, as defined in the preceding section and let us multiply the polynomial $C_{\frac{n-1}{2}}(x)$ in this expression by a linear factor $1 - ax$ to get a polynomial $C_p(x) = C_{\frac{n}{2}}(x)$ of correct degree. Thus we have

$$(5.2) \qquad C_p(x) = (1 - ax)T_{p-1}(1 - 2x), \; p = \frac{n}{2},$$

$$(5.3) \qquad S_q(x) = \frac{1 - x}{q} U_{q-1}(1 - 2x) \quad , \; q = p-1.$$

In order to maximize $B(n,x,\vec{\alpha})$ for $x \sim 0$ we must choose $a = 2$. We then obtain

$$(5.4) \qquad b(n,\vec{\alpha}) = (n-2)^2 \min_{0\le x\le 1} \left[\frac{1}{1 - x} + 4x \operatorname{cotg}^2\left[\frac{n-2}{2} \arccos(1-2x)\right]\right].$$

In figure 5.1 the behaviour of the functions $1/1-x$ and $x \operatorname{cotg}^2[\frac{1}{2}(n-2) \arccos(1-2x)]$ is illustrated.

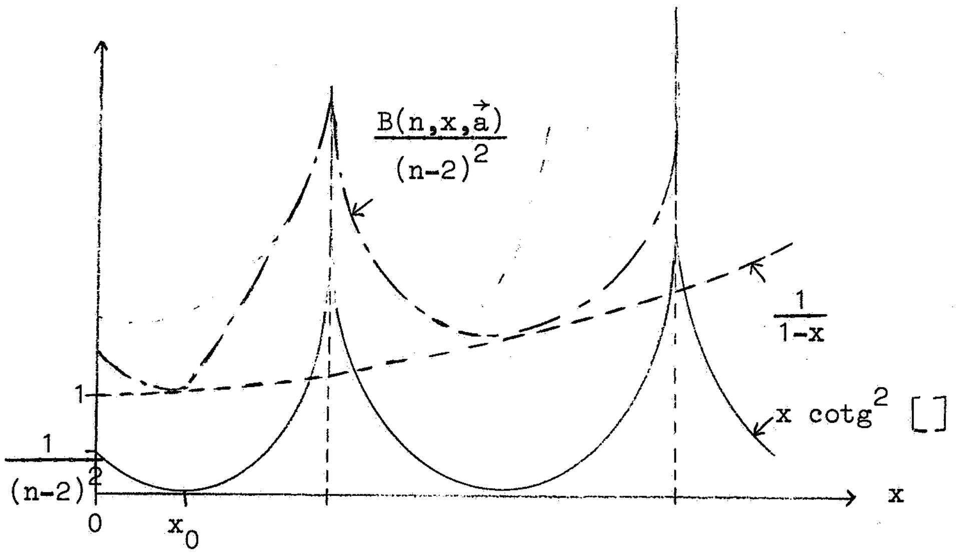


fig. 5.1

From this figure it is clear that the minimum of $B(n,x,\vec{\alpha})$ is reached in the interval $[0,x_0]$, where x_0 is the first positive zero of cotg [ ], i.e.

(5.5) $$x_0 = \frac{1}{2}\,(1 - \cos\frac{\pi}{n-2}).$$

Let us take $n \geq 6$. Then x_0 is relatively small, so that we may approximate the function $B(n,x,\vec{\alpha})$ on the interval $[0,x_0]$ by

(5.6) $$B(n,x,\vec{\alpha}) \sim (n-2)^2\,\{\frac{1}{1-x} + 4x\,\text{cotg}^2(n-2)\sqrt{x}\}.$$

The first derivative of $B(n,x,\vec{\alpha})$ is given by

(5.7) $$\frac{d}{dx}\,B(n,x,\vec{\alpha}) \sim \frac{(n-2)^2}{\sin^2 y}\,\{\sin^2 y + 4\cos^2 y - 4y\,\text{cotg}\,y\},$$

where $y = (n-2)\sqrt{x}$.
From this expression it can easily be derived that $B(n,x,\vec{\alpha})$ reaches its minimum for

(5.8) $$y \sim 1.36 \quad \text{or} \quad x \sim \frac{1.85}{(n-2)^2},$$

so that

$$(5.9) \qquad b(n,\vec{\alpha}) \sim \frac{(n-2)^4}{(n-2)^2 - 1.85} + .296.$$

For instance, we have $b(6,\vec{\alpha}) \sim 18.4$ and $b(8,\vec{\alpha}) \sim 38.1$. These values lead to effective time steps which are considerably smaller than the effective time steps obtained for odd values of n. Therefore, the operator defined by (5.2) and (5.3) is not recommended in actual computation.

Of course, one may look for more appropriate polynomials of even degree, for instance by employing numerical methods. However, it may be remarked, there are reasons to believe that such an approach will be very difficult: Let us consider the function $b(n,\vec{\alpha})$ where n is odd. From section 4 it may be concluded that the optimal point in the $\vec{\alpha}$-space, also is a singular point of the function $b(n,\vec{\alpha})$. Further, $b(n,\vec{\alpha})$ will have many local maxima separated from each other by lines of discontinuity. It is plausible to suppose that this will also be the case when n is even. Hence, such difficulties have to be overcome when using numerical methods.

6. Extension to non-imaginary eigenvalues

In this section we allow the operator D_0 to also have eigenvalues in the negative halfplane of the complex δ-plane. This extented operator will be denoted by D; the largest value of the moduli of the purely imaginary eigenvalues will still be denoted by $\sigma(D_0)$.

Let us consider the operator $P_n(D)$ for odd values of n as defined in the preceding sections and let us determine the domain in the δ-plane where the polynomial $P_n(\delta)$ has values within or on the unit circle, i.e. the points δ which satisfy the inequality

(6.1) $$|P_n(\delta)| = |1 + \tau\delta + \beta_2\tau^2\delta^2 + \ldots + \beta_n\tau^n\delta^n| \leq 1.$$

We shall put $\tau = \sqrt{b}/\sigma(D_0) = (n-1)/\sigma(D_0)$.
Further, it is convenient to introduce the new variable

(6.2) $$\zeta = \frac{\delta}{\sigma(D_0)} .$$

We then obtain

(6.2') $$P_n(\delta) = \tilde{P}_n(\zeta) \equiv 1 + (n-1)\zeta + \alpha_2\zeta^2 + \alpha_3(n-1)\zeta^3 + \alpha_4\zeta^4 + + \ldots + \alpha_n(n-1)\zeta^n \equiv C_p(-\zeta^2) + (n-1)\zeta S_p(-\zeta^2),$$

where the coefficients $\alpha_2, \ldots, \alpha_n$ and the polynomials C_p and S_p are defined as in sections 3 and 4.
In figure 6.1 the curves $|\tilde{P}_n(\zeta)| = 1$ are given for n = 2, 3, 4, 5, 7, 9.

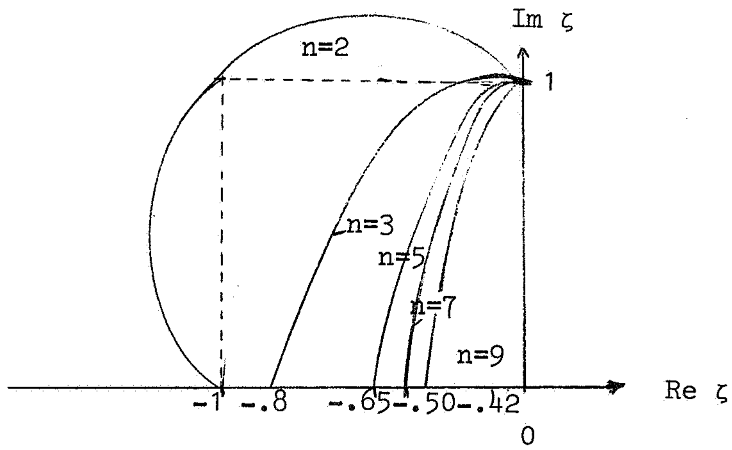


fig. 6.1 Contours $|P_n(\zeta)| = 1$

Since $|\tilde{P}_n(\zeta)| = |\tilde{P}_n(\overline{\zeta})|$ we have omitted the part of the contour for which $\text{Im}\ \zeta < 0$.

For $n \to \infty$ the domain $|\tilde{P}_n(\zeta)| \leq 1$ converges to the segment $[-i,i]$ of the imaginary axis. In connection with this it may be remarked that for $n = \infty$ we have

$$(6.3) \qquad |P_\infty(\delta)| = |1 + \tau\delta + \frac{1}{2!}\tau^2\delta^2 + \ldots + \frac{1}{m!}\tau^m\delta^m + \ldots|$$

$$= |\exp \tau\delta| = \exp(\tau\ \text{Re}\ \delta),$$

so that for $\text{Re}\delta \leq 0$ $|P_n(\delta)| \leq 1$. Hence for $n = \infty$ the eigenvalues δ are allowed to be in the whole negative half-plane.

In practice, however, we use finite values for n and we have a finite domain in the δ-plane in which the eigenvalues of D are admitted. For instance, of practical importance is the situation where a linear friction term is added to the difference scheme (1.4), i.e.

$$(6.4) \qquad u_{k+1} = (1 + \tau_k D_0 - \lambda\tau_k)u_k,$$

where λ is a coefficient of friction. An example of such a scheme is given in [2] p. 65. The eigenvalues δ of the operator

$$(6.5) \qquad D = D_0 - \lambda$$

are situated on the line-segment $[-\lambda - i\sigma(D_0), -\lambda + i\sigma(D_0)]$ of the δ-plane or the line-segment $[-\lambda/\sigma(D_0) - i, -\lambda/\sigma(D_0) + i]$ of the ζ-plane.

In general, $\sigma(D_0)$ is large, hence the eigenvalues are situated on a line close to the imaginary axis. In table 6.1 the upper bounds for the coefficient of friction λ are given for some values of n.

Table 6.1

n	2	3	4	5	7	9
λ	1.00 $\sigma(D_0)$	.39 $\sigma(D_0)$	.21 $\sigma(D_0)$	.08 $\sigma(D_0)$	.025 $\sigma(D_0)$	.01 $\sigma(D_0)$

It is interesting that for a given value of n and λ one obtains always a stable scheme by choosing a finer grid, i.e. by making $\sigma(D_0)$ larger.

References

[1] Handbook of Mathematical functions, National Bureau of Standards, Applied Mathematics Series 55, edited by M. Abromowitz and J.A. Stegun (1964).

[2] Houwen, P.J. van der, Finite difference methods for solving partial differential equations, MC Tract 20, Mathematisch Centrum, Amsterdam (1968).

[3] Lauwerier, H.A., Damsté, B.R., A numerical treatment. The North Sea Problem VIII, Proc. Kon. Ned. Ak. v. Wetensch. A66, 167-184 (1963).