

## NUMERICAL STABILIZATION OF RECURRENCE RELATIONS WITH VANISHING SOLUTIONS

H.F. ARNOLDUS

*Fysisch Laboratorium, Rijksuniversiteit Utrecht, Postbus 80 000, 3508 TA Utrecht, The Netherlands*

Received 10 May 1984

Homogeneous recurrence relations exhibit a highly numerical unstable behaviour in step-by-step evaluation of successive terms. It is pointed out that this is a result of the presence of vanishing solutions, which are always added to initial values for the recursion scheme, due to finite machine accuracy. Stabilization of the recursion is shown to be identical with resolving these vanishing contributions with sufficient accuracy. To this end, explicit analytical expressions for these solutions, as products of continued fractions, are given. Application of these vanishing solutions enables us to construct the self-consistent, numerical stable general solution of the recursion relation.

### 1. Introduction

Quantum mechanical inelastic scattering problems involve an immense number of radial partial wave function matrix elements, which differ only by the quantum number of orbital angular momentum. Especially in heavy ion collisions, these matrix elements are very cumbersome and even with present-day high-speed computers very hard to calculate [1-5]. Fortunately, many of these matrix elements are connected by recurrence relations, so only a few have to be found by explicit numerical integration [6,7]. It has, however, been taken for granted that the successive generation of integrals from some initial ones is limited because the relations are highly unstable. A similar problem appears in statistical mechanics on lattices, where the values of microscopic thermodynamic functions on the lattice points are connected by similar relations. With very general arguments, it can be shown that a recurrence relation has exponentially increasing solutions, except one. The point at issue in statistical mechanics is to find this unique decreasing solution, but due to the intrinsic numerical instability of the recurrence relations, only crude approximations can be found, which match to an asymptotic vanishing solution.

In this paper I will point out how these two

problems are related. The instability of the recurrence relations is a reflection of the feature that the vanishing solution is always eclipsed by the exponentially increasing component of the general solution. In order to solve this apparent problem, I give an explicit analytical expression for the vanishing solution as a product of continued fractions, which can be evaluated numerically in a stable fashion, and I show how adequate use of this solution stabilizes recursion schemes. The method described in this paper can be applied very generally, and I will illustrate common features with a specific example from heavy ion scattering theory. The computer programs were run on our CDC 175/100 and I used double precision variables with a machine accuracy of 28 figures.

### 2. Instability caused by the vanishing solutions

Consider the homogeneous three-term recurrence relation

$$\rho_{k+1}X_k - \xi_{k+1}X_{k+1} + \gamma_{k+1}X_{k+2} = 0, \quad (1)$$
$$k = 0, 1, 2, \dots$$

with arbitrary coefficient functions  $\rho_k$ ,  $\xi_k$  and  $\gamma_k$ .

If we know two neighbouring initial values  $X_i$ ,  $X_{i+1}$ , we can find every other  $X_k$  with step-by-step up- and downward recursion with (1). This solution involves the two arbitrary constants  $X_i$  and  $X_{i+1}$  but the general solution is a linear combination of two linearly independent solutions which might be generated for instance with the initial values  $X_0 = 0$ ,  $X_1 = 1$  and  $X_0 = 1$ ,  $X_1 = 0$ , respectively. Then any two given values  $X_i$ ,  $X_j$  with  $i \neq j$  determine uniquely the linear combination.

This seems all straight forward, but in order to find out what problems appear, suppose that eq. (1) reduces approximately to

$$X_k - \beta X_{k+1} + X_{k+2} = 0, \quad \beta > 1 \quad (2)$$

for large  $k$  values, as is the case in many practical situations. If we start the recursion with arbitrary  $X_0$  and  $X_1$ , it is easy to see that  $X_k \sim \beta^k$ , so the solutions will increase exponentially with  $k$ , but since equations of type (2) are recurrence relations for hypergeometric functions, we know that there also exists a solution which tends to zero for  $k \rightarrow \infty$ . Any linear combination of this vanishing solution with an exponentially increasing solution is again exponentially increasing, so there exists only one vanishing solution for  $k \rightarrow \infty$ . Instead of searching for solutions, determined by initial values  $X_i$  and  $X_j$ , we can also try to find this particular solution, which is now determined by for instance  $X_0 = 1$  and the requirement that  $X_k \rightarrow 0$  for  $k \rightarrow \infty$ . If we have two linearly independent solutions, we can try to find the linear combination, which gives this special solution. I intend to explain that this procedure is numerically not feasible, at least if we want to find  $X_0$ ,  $X_1$ , ...,  $X_k$ , ... also for large  $k$  values. Suppose we have the independent solutions up to  $k = 50$ . Then  $\beta^k$  is already  $10^{15}$  for  $\beta \approx 2$ , so if a linear combination of two of these  $X_k$ 's should yield a number which is smaller than one, both terms must be equal to 15 figures. Subtracting these two numbers gives indeed a small number, but with an accuracy that is at least 15 figures less than the accuracy of the  $X_k$ 's and if we increase  $k$ , no figure will be left so the  $X_k$ 's become completely random. I will now point out how this problem is related to the numerical instability of relation (1).

Let me first show by an example how severe the

problem of instability actually can be. In the next sections a closed relation for the vanishing solution is obtained, which can be evaluated with almost machine accuracy. For a specific example, some  $X_k$ 's are given in the first column of table 1. To see what happens in step-by-step recursion, take  $X_0$  and  $X_1$  from this solution and calculate  $X_2$ ,  $X_3$ , ... with eq. (2). These results are printed in the third column of table 1 upto  $k = 1000$ . As far as  $k \approx 100$ , the results are reasonable but for  $k \geq 150$  the  $X_k$ 's start to increase and for  $k = 1000$ , there is a factor  $10^{209}$  discrepancy with the desired result and even the sign is wrong. Note that both in column 1 and column 3, every set of three subsequent  $X_k$ 's obeys within machine accuracy relation (1), so to 28 figures. The only approximation that has been made is that the initial  $X_0$  and  $X_1$  are given within machine accuracy. This deviation from the exact value gives effectively that an originally very small amount of an increasing component of the solution is added. After many steps, this small contribution has grown exponentially with  $\beta^k$  and so it overwhelms completely the vanishing solution. If we had started the up-recursion with values for  $X_0$  and  $X_1$ , truncated after for instance 25 figures, a completely different column 3 would have resulted. That is, the result after  $k \approx 140$  is random or the up-recursion is unstable.

It is illuminating to make these considerations more explicit. To this end we introduce a second vanishing solution, defined by  $X_0 = 0$  and just as the former one, this solution is determined up to an arbitrary overall constant. Let us denote these solutions by  $F_k$  and  $G_k$  with

$$\begin{aligned} F_k &\rightarrow 0 \quad \text{if } k \rightarrow \infty, \\ G_0 &= 0. \end{aligned} \quad (3)$$

If we exclude the trivial solution  $X_k = 0$  for all  $k$ , it can be shown that  $F_0 \neq 0$  (see later on) so  $F_k$  and  $G_k$  are independent solutions and we can write the linear combination as

$$X_k = aF_k + bG_k \quad (4)$$

with  $a$  and  $b$  independent of  $k$ . Since there is only one decreasing solution for  $k \rightarrow \infty$ ,  $G_k$  is exponentially increasing. Now we can return to the problem of instability. If we take  $X_0 = F_0$  and  $X_1 = F_1$

Table 1  
Instability in finding the vanishing solutions of a three-term homogeneous recurrence relation

$k$	$F_k$	$G_k$	$UP_k$
0	0.1000000000D+001	0.0000000000D+000	0.1000000000D+001
1	0.1058578352D+001	0.9327299328D-116	0.1058578352D+001
2	0.1114502599D+001	0.1859872539D-115	0.1114502599D+001
3	0.1166787826D+001	0.2776458694D-115	0.1166787826D+001
4	0.1214436295D+001	0.3676438919D-115	0.1214436295D+001
5	0.1256469935D+001	0.4552884248D-115	0.1256469935D+001
6	0.1291965748D+001	0.5398240714D-115	0.1291965748D+001
7	0.1320092264D+001	0.6204630081D-115	0.1320092264D+001
8	0.1340144862D+001	0.6964184595D-115	0.1340144862D+001
⋮			
153	0.4845757450D-013	0.1696611305D-103	0.4275051473D-013
154	0.3674514560D-013	0.2208379507D-103	0.2931550035D-013
155	0.2785935542D-013	0.2875207872D-103	0.1818773344D-013
156	0.2111913907D-013	0.3744270628D-103	0.8524164474D-014
157	0.1600724690D-013	0.4877148154D-103	-0.3985002197D-015
158	0.1213091527D-013	0.6354241733D-103	-0.9243478078D-014
159	0.9191960753D-014	0.8280545597D-103	-0.1866213201D-013
160	0.6964044986D-014	0.1079319480D-102	-0.2934209520D-013
⋮			
995	0.7649855340D-118	0.2395165191D+000	-0.8056854786D+089
996	0.5468918175D-118	0.3187619340D+000	-0.1072251143D+090
997	0.3744912865D-118	0.4242270912D+000	-0.1427014756D+090
998	0.2334258592D-118	0.5645875630D+000	-0.1899159200D+090
999	0.1119591533D-118	0.7513895915D+000	-0.2527523717D+090
1000	0.0000000000D+000	0.1000000000D+001	

we obviously have  $b = 0$ , but due to the finite machine accuracy, we effectively have  $b \approx 0$ . Since this  $b$  is multiplied by the increasing  $G_k$  in (4), we will always find that the second term on the right hand side of (4) dominates over  $aF_k$  for  $k$  sufficiently large. In other words, the instability results from the fact that the vanishing solution  $G_k$  cannot be resolved properly for small  $k$ , which is the inverse problem of resolving the  $F_k$  from an increasing solution as indicated in the first part of this section. The impossibility of finding  $F_k$  for  $k$  large is identical to the instability problem, i.e. the impossibility of finding  $G_k$  in  $X_k$  for  $k$  small with sufficient accuracy.

There is another correspondence between the two vanishing solutions. The  $F_k$  is the solution of (1) which tends to zero for  $k \rightarrow \infty$  if we start the upward recursion with  $X_0 = F_0$  and  $X_1 = F_1$ . In precisely the same way,  $G_k$  is a solution of (1) in downward recursion if we start with some  $X_n = G_n$  and  $X_{n+1} = G_{n+1}$ . Since  $G_0 = 0$  and  $F_0 \neq 0$ , this is

the only solution that vanishes for  $k \rightarrow 0$  in downward recursion. The  $G_k$  decreases as  $\beta^{k-n}$  for  $k \rightarrow 0$ , just as  $F_k$  decreases with  $\beta^{-k}$  for  $k \rightarrow \infty$ , so the behaviour is similar. Just as  $F_k$  for  $k$  large, cannot be found from upward recursion, the  $G_k$  for  $k \rightarrow 0$  cannot be found from downward recursion. Furthermore, the mixture of a small amount of  $G_k$  in  $X_0, X_1$  can make an upward recursion unstable and the mixture of a small amount of  $F_k$  might destroy downward recursion.

### 3. The self-consistent solution

Suppose we are interested in  $X_k$  with  $k = 0, 1, 2, \dots, k_{\max}$ , with  $k_{\max}$  large, then it will be clear that we have to abandon the step-by-step method. In this section I will show that explicit application of the vanishing solutions offers a possibility to stabilize the recursion. To achieve these special solutions, write the homogeneous eq. (1) as the

inhomogeneous tridiagonal matrix equation

$$\begin{pmatrix} -\xi_1 & \gamma_1 & & & & \\ \rho_2 & -\xi_2 & & \gamma_2 & & 0 \\ & & \ddots & & & \\ & & & \rho_k & -\xi_k & \gamma_k \\ 0 & & & & \ddots & \\ & & & & & \rho_n & -\xi_n \end{pmatrix} \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} -\rho_1 X_0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ -\gamma_n X_{k_{\max}} \end{pmatrix}. \quad (5)$$

where  $n = k_{\max} - 1$ . Now we can introduce solutions of this set which vanish at the boundaries  $k = 0$  and  $k = k_{\max}$  and are normalized by

$$\begin{aligned} F_0 &= 1, & F_{k_{\max}} &= 0 \\ G_0 &= 0, & G_{k_{\max}} &= 1. \end{aligned} \quad (6)$$

Note that, due to the truncation,  $F_k$  is an approximation of the solution that vanishes exactly for  $k \rightarrow \infty$ , but the introduction of  $k_{\max}$  is not an approximation of the equations.

The matrix equation (5) with boundary conditions (6) can be solved explicitly, with result

$$\begin{aligned} F_k &= (-1)^k \frac{\rho_k \cdots \rho_2 \rho_1}{u_k \cdots u_2 u_1}, \\ G_k &= (-1)^{n-k+1} \frac{\gamma_k \cdots \gamma_{n-1} \gamma_n}{v_k \cdots v_{n-1} v_n}, \\ &k = 1, 2, \dots, n, \end{aligned} \quad (7)$$

where the  $u_k$  and  $v_k$  follow from

$$\begin{aligned} u_n &= -\xi_n, & u_k &= -\xi_k - \frac{\gamma_k \rho_{k+1}}{u_{k+1}}, \\ v_1 &= -\xi_1, & v_{k+1} &= -\xi_{k+1} - \frac{\gamma_k \rho_{k+1}}{v_k}, \\ && &k = 1, 2, \dots, n-1, \end{aligned} \quad (8)$$

as can be checked by inspection. Every set  $F_k$ ,  $F_{k+1}$ ,  $F_{k+2}$  and  $G_k$ ,  $G_{k+1}$ ,  $G_{k+2}$  fulfils the recursion relation with machine accuracy, if evaluated

with (7) and (8), and at the upper and lower bounds of  $k$ , the solutions are exact. So these are self-consistent vanishing solutions, which means that for  $k \rightarrow k_{\max}$  and  $k \rightarrow 0$ , respectively, these solutions are forced to become zero. A blow up, due to instabilities, is excluded.

Numerical evaluation of (7) and (8) is trivial. In table 1 we printed some  $F_k$  and  $G_k$ , where we have chosen an example from nuclear scattering theory. In that case, the  $X_k$ 's are Coulomb integrals and the coefficient functions are defined by [7]

$$D_k(\eta) = \left\{ 1 + \left( \frac{\eta}{k+1} \right)^2 \right\}^{1/2}, \quad (9)$$

$$\rho_{k+1} = \frac{k+1}{k+2} D_{k+1}(\eta) D_k(\eta'), \quad (10)$$

$$\begin{aligned} \xi_{k+1} &= \frac{\eta'}{\eta} \frac{k+3/2}{k+2} D_{k+1}^2(\eta) \\ &+ \frac{\eta}{\eta'} \frac{k+5/2}{k+2} D_{k+1}^2(\eta'), \end{aligned} \quad (11)$$

$$\gamma_{k+1} = \frac{k+3}{k+2} D_{k+2}(\eta) D_{k+1}(\eta'), \quad (12)$$

with  $\eta > 0$ ,  $\eta' > 0$ . The factor  $\beta$  from eq. (2), which implies the increasing behaviour, is obviously

$$\beta = \frac{\eta}{\eta'} + \frac{\eta'}{\eta} \geq 2. \quad (13)$$

For the results in table 1 we used  $\eta = 30$ ,  $\eta' = 40$ , so  $\beta \approx 2.08$ , and we took  $k_{\max} = 1000$ . We see that  $F_k$  indeed tends to zero very smoothly and since  $F_{999}$  is already  $10^{-119}$ , it can be assumed that this  $F_k$  is an excellent approximation for the exact vanishing solution, with  $k_{\max} = \infty$ . Furthermore, we find that, with the normalisation  $G_{1000} = 1$ , the  $G_k$  for small  $k$  is very small indeed ( $\approx 10^{-116}$ ), compared with the machine accuracy in  $F_1$  which is about  $10^{-28}$ , so this  $G_k$  component could never be resolved in anyway if added to  $F_1$ . If we use  $F_0$ ,  $F_1$ , found in this way, for upward recursion, as has been done in column 3 of table 1, this implies that after 998 recursions a  $G_1$  deviation of  $10^{-116}$  in  $F_1$ , gives rise to a contribution  $G_{999} \approx 1$ , which is a factor of  $10^{119}$  larger than the desired component  $F_{999}$ . This illustrates explicitly that  $F_k$  for  $k$  large can never be found by upward recursion.

The general solution can now be represented by the linear combination

$$X_k = aF_k + bG_k, \quad k = 0, 1, 2, \dots, k_{\max}, \quad (14)$$

where  $a$  and  $b$  can be matched with two initial values  $X_i$  and  $X_j$ . From (6) we have the obvious possibility

$$a = X_0, \quad b = X_{k_{\max}} \quad (15)$$

and I will argue that this choice stabilizes the recursion for every possible  $X_0$  and  $X_{k_{\max}}$ . If we determine  $a$  and  $b$  in this way with some initial value routine, we have by the same reasoning as previously, a self-consistent solution, i.e. the boundary values are exact and every triple  $X_k, X_{k+1}, X_{k+2}$  obeys the recurrence relation with the same accuracy as  $X_0$  and  $X_{k_{\max}}$ , so no figures are lost due to the recursion. There is however a more profound reason why the matching should be performed with (15). It might seem that also, for instance

$$b = \frac{X_1 - X_0 F_1}{G_1} = a \frac{X_1/X_0 - F_1}{G_1} \quad (16)$$

is suitable. It is indeed analytically correct, but it does not stabilize the recursion, because if  $a$  and  $b$  are of the same order of magnitude, the numerator  $X_1/X_0 - F_1$  should be of the order of  $G_1$ , which is very small in comparison with  $F_1$ . This implies that  $F_1$  should equal  $X_1/X_0$  to many figures and the subtraction yields again random results. This is an equivalent formulation of the explanation of the instability caused by the vanishing solution  $G_k$ , so the choice (16) should consequently be rejected. That combination of the application of the vanishing solutions and the matching (15) stabilizes the recursion, can also be seen from a different point of view. If we take  $k = 1$  in (14), we see that  $bG_1$  is negligible with respect to the first term  $aF_1$ , so the  $b$  cannot be resolved from  $X_1$ . Similarly, we cannot find  $a$  from  $X_k$  with  $k$  large. Now it becomes clear why the use of the two vanishing solutions, in combination with (15), actually solves the problem. The coefficients  $a$  and  $b$  can be found very precisely due to the disappearance of one of the components at the matching points  $k = 0$  and  $k = k_{\max}$ . These  $a$  and  $b$ , however, also determine with the same precision

the contribution of  $F_k$  at large  $k$  values and the contribution of  $G_k$  for small  $k$  values, so the  $F_k$  is even very accurately resolved, with respect to  $G_k$ , for  $k$  large and, vice versa, we have the  $G_k$  contribution for small  $k$  with the same accuracy as the  $F_k$  component, which is many orders of magnitude larger. This is not possible for any other set of independent solutions or with step-by-step recursion. A special case is  $b = 0$ , which yields the vanishing solution  $F_k$  up to arbitrary large  $k$  values. The distinction with upward recurrence is that now  $b$  is identically zero, and not zero within machine accuracy.

The choice  $k = 0$  and  $k = k_{\max}$  as limiting values is arbitrary. We could equally well take  $k_i$  and  $k_j$ , but then (15) gives  $a = X_{k_i}, b = X_{k_j}$ . With (14) we then find  $X_k = 0$  for all  $k$  if these two initial values are zero. This proves that either all  $X_k$ 's are zero or not more than one, as mentioned earlier.

#### 4. Influence of the truncation at $k_{\max}$

The vanishing solutions  $F_k$  and  $G_k$  are products of continued fractions, as follows from (7) and (8).

$$\begin{aligned} \frac{F_k}{F_{k-1}} &= \frac{-\rho_k}{u_k} \\ &= \frac{-\rho_k}{-\xi_k - \frac{\gamma_k \rho_{k+1}}{-\xi_{k+1} - \frac{\gamma_{k+1} \rho_{k+2}}{-\xi_{k+2} - \dots - \xi_n}}} \end{aligned} \quad (17)$$

$$\begin{aligned} \frac{G_k}{G_{k+1}} &= \frac{-\gamma_k}{v_k} \\ &= \frac{-\gamma_k}{-\xi_k - \frac{\gamma_{k-1} \rho_k}{-\xi_{k-1} - \frac{\gamma_{k-2} \rho_{k-1}}{-\xi_{k-2} - \dots - \xi_1}}} \end{aligned} \quad (18)$$

The solution  $F_k$  is important in its own right in the

Table 2  
Example of truncation at  $k_{\max}$

$k_{\max}$	$F_1$
10	0.9066631950D+00
20	0.9774900333D+00
30	0.1027226473D+01
40	0.1054550649D+01
50	0.1058408539D+01
60	0.1058574684D+01
70	0.1058578300D+01
80	0.1058578351D+01
90	0.1058578352D+01
100	0.1058578352D+01
⋮	
1000	0.1058578352D+01

limit  $k_{\max} \rightarrow \infty$ , whereas the  $G_k$  then vanishes for all  $k$ . Note that  $k_{\max} \rightarrow \infty$  in (17) simply means that the continued fraction is not terminated anymore after  $n$  steps. The values of  $F_k$  for small  $k$  are not expected to depend on  $k_{\max}$  any more if  $k_{\max}$  is sufficiently large. This is shown by an example in table 2, where the continued fraction  $F_1 = -\rho_1/u_1$  is printed for different truncations  $k_{\max}$  and we observe that the convergence is very fast. At  $k_{\max} = 90$ , we already have an accuracy of ten figures in  $F_1$ . I recall once more that the truncation is no approximation at all in the case where we solve the recursion for given  $X_0$  and  $X_{k_{\max}}$ .

## 5. Conclusions

In this paper I pointed out that the problem of finding the vanishing solution of a three-term homogeneous recurrence relation is related to the intrinsic numerical instability of that relation. Exponentially increasing components will always obscure the vanishing solution and after a number of upward recursion steps, this solution cannot be resolved any more. The connection with the instability of the system is that in downward recursion this initially eclipsed component can become much larger than any wanted solution. Since finite mac-

hine accuracy is always present, giving rise to random mixing of increasing components, step-by-step recursion is limited to a small number of steps (20–50) and an explicit construction of the vanishing solution, which might be physically important, terminates rather fast. In sections 3 and 4, I presented explicit expressions for the two vanishing solutions as products of continued fractions. These expressions were proven to be self-consistent, which means that any mixing with increasing components is automatically avoided. Numerical evaluation of the continued fractions is almost trivial and numerically stable. It was furthermore pointed out how proper use of these vanishing solutions stabilizes the general solution of the recurrence relation, where the number of steps  $k_{\max}$  is allowed to be arbitrary large. This was illustrated with a numerical example from a DWBA scattering problem, where we took  $k_{\max} = 1000$ .

## Acknowledgements

A discussion with T. Nieuwenhuizen from Utrecht University, where he introduced me to the problem of the vanishing solution, is cordially granted and I greatly acknowledge the spontaneous discussions with L.D. Tolsma from Eindhoven University of Technology on a related problem.

## References

- [1] L.C. Biedenharn, J.L. McHale and R.M. Thaler, Phys. Rev. 100 (1955) 376.
- [2] K. Alder, A. Bohr, T. Huus, B. Mottelson and A. Winter, Rev. Mod. Phys. 28 (1956) 432.
- [3] M. Samuel and U. Smilansky, Comput. Phys. Commun. 2 (1971) 455.
- [4] G.H. Rawitscher and C.H. Rasmussen, Comput. Phys. Commun. 11 (1976) 183.
- [5] H.F. Arnoldus, Comput. Phys. Commun. 32 (1984) 421.
- [6] L.E. Wright, Phys. Rev. C20 (1979) 393.
- [7] J. Raynal, Phys. Rev. C23 (1981) 2571.