Journal of Global Optimization **6:** 293–311, 1995.

Simple Global Minimization Algorithm for One-Variable Rational Functions

JAROSLAW KOSTROWICKI and HAROLD A. SCHERAGA
Baker Laboratory of Chemistry, Cornell University, Ithaca, NY 14853-1301, U.S.A.

(Received: 19 April 1994; accepted: 10 January 1995)

Abstract. Sturm's chain technique for evaluation of a number of real roots of polynomials is applied to construct a simple algorithm for global optimization of polynomials or generally for rational functions of finite global minimal value. The method can be applied both to find the global minimum in an interval or without any constraints. It is shown how to use the method to minimize globally a truncated Fourier series. The results of numerical tests are presented and discussed. The cost of the method scales as the square of the degree of the polynomial.

Key words: Global optimization, univariate optimization, polynomials, rational functions, Sturm's chain

1. Introduction

Several algorithms have been devised to address the global minimization problem for functions of one variable. Some of them are of general nature (Cetin et al., 1993; Hansen, 1979), while others are restricted to certain classes of functions as, for example, polynomials (Visweswaran and Floudas, 1992; Wingo, 1985). A large class of algorithms studies values of a function and/or its derivatives at given points and then moves according to the local properties found in order to arrive at points with lower function values (e.g., Cetin *et al.*, 1993). Other methods construct globally valid lower bounds for the minimized function and proceed by refining these bounds (e.g., Visweswaran and Floudas, 1992).

The method proposed here is closer in spirit to the second group of methods. To find the globally minimal value of the objective function $f(x)$, we examine the existence of real zeros for the shifted function: $f(x) + \gamma$ for several values of γ. The procedure is terminated, with given accuracy ϵ, when the globally minimal value of $f(x) + \gamma$ is sufficiently close to zero i.e. the function $f(x) + \gamma + \epsilon$ has no real zeros whereas $f(x) + \gamma - \epsilon$ does have at least one of them. Finally we locate the position where f actually assumes this globally minimal value. The method is designed for a function of the form: $f(x) = W(x)/V(x)$ where both W and V are polynomials. It is also assumed that the global minimum exists, which means that for W and V not having common divisors, all real zeros of V (if any) are of even multiplicity, and the sign of V in their vicinity is the same as the sign of W. This implies that $V(x)$ has the same sign for all values of x. For simplicity we assume it to be $+$. Because the limit of f in $\pm\infty$ should not be $-\infty$, the leading term of V should have equal or greater degree than the leading term of W; or, if the leading term of V has lower degree than the leading term of W, then the leading coefficients should have the same sign and the leading exponents should differ by an even number. Of course, for $V(x) = 1$, the algorithm serves as a global minimizer for polynomials.

Since $f_\gamma(x) \equiv f(x) + \gamma = [W(x) + \gamma V(x)]/V(x)$ has zeros if and only if $W(x) + \gamma V(x)$ has zeros, one needs a test for the existence of real roots for polynomials. This can be achieved by Sturm's chain technique (van der Waerden, 1991). The parameter γ is then adjusted until $f_\gamma(x)$ has its globally minimal value equal to zero. The position of the minimum can be derived from the last polynomial in the Sturm chain for this γ value.

2. The Algorithm

For convenience, we include Sturm's theorem here (see for example van der Waerden, 1991, for the proof). First, we define Sturm's chain. For a given polynomial X_0, one forms the sequence of polynomials $X_1, \ldots, X_r$ such that $X_1 = X_0'$ (the derivative of X_0); then, for all $i \geq 2$, the polynomial $-X_i$ is the remainder after dividing X_{i-2} by X_{i-1} (Euclidean algorithm):

$$X_{i-2} = Q_{i-1}X_{i-1} - X_i \tag{1}$$

The divisions are continued until $X_{r-1} = Q_r X_r$. The sequence $X_1, \ldots, X_r$ is called Sturm's chain for X_0. It is usually convenient to divide every element of Sturm's chain by the absolute value of its leading coefficient before proceeding to the next element. For a given number a, which is not a root of X_0, let $w(a)$ be a number of sign changes in the sequence $X_0(a), \ldots, X_r(a)$. While counting sign changes, one should disregard zeros. Hence, for example, for the sequence: $3, 2, -1, 0, 5, 6, 0, -2$ the number of sign changes is 3. We will also need the limit values $w(\pm\infty) = \lim_{a\to\pm\infty} w(a)$. They can easily be evaluated from the leading coefficients of the polynomials X_i. The sign of X_i for $a \to +\infty$ is just the sign of the leading coefficient α_{im_i}, whereas the sign for $a \to -\infty$ is that of $(-1)^{m_i}\alpha_{im_i}$, where $X_i(x) = \sum_{n=0}^{m_i} \alpha_{in} x^n$

Sturm's theorem is:

If $b < c$ (b may be $-\infty$ and c may be ∞) and $X_0(b) \neq 0$, $X_0(c) \neq 0$, then the number of distinct real roots of X_0 between b and c is equal to $w(b) - w(c)$ (multiple roots are counted once).

For example:

$$X_0(x) = x^4 - 11x^3 + 41x^2 - 61x + 30$$

$$X_1(x) = x^3 - \frac{33}{4}x^2 + \frac{41}{2}x - \frac{61}{4}$$

$$X_2(x) = x^2 - \frac{34}{7}x + \frac{191}{35}$$

$$X_3(x) = x - \frac{25}{11}$$

$$X_4(x) = 1.$$

One can easily check that $w(-\infty) = 4$ and $w(\infty) = 0$. Hence according to Sturm's theorem X_0 has 4 real roots and indeed $X_0(x) = (x-1)(x-2)(x-3)(x-5)$.

The method proposed in this paper finds the global minimum of a rational function f in an interval $[b, c]$ (b may be $-\infty$ and c may be ∞) by examining the existence of zeros of the shifted function f_γ. Let us denote by $\gamma^\diamond$ the exact border value of γ, *i.e.* such that for every $\sigma > 0$ the function $f_{\gamma^\diamond+\sigma}$ has no roots and also that there exists $\delta > 0$ that for every σ, $-\delta < \sigma \leq 0$ the function $f_{\gamma^\diamond+\sigma}$ does have roots. For a given accuracy ϵ, we find a lower bound $\underline{\gamma}$ and an upper bound $\overline{\gamma}$ of the exact value $\gamma^\diamond$ ($\underline{\gamma} < \gamma^\diamond < \overline{\gamma}$) such that $\overline{\gamma} - \underline{\gamma} < \epsilon$.

The initial bounds are constructed in the following way. In the first case, *i.e.*, when f has no zero and is positive in $[b, c]$, we take $\overline{\gamma} = 0$ and $\underline{\gamma} = -f(x_0)$ where x_0 is an arbitrary point in $[b, c]$. In the second case, *i.e.*, when f has no zero and is negative, then $\underline{\gamma} = -f(x_0)$ and $\overline{\gamma} = -f(x_0) + \Delta$, where Δ is chosen as a sufficiently large number so that there is no zero for $f_{\overline{\gamma}}$. The last case, *i.e.*, when f has zeros in $[b, c]$, we set $\underline{\gamma} = 0$ and $\overline{\gamma} = \Delta$ where Δ is chosen, again, so that $f_{\overline{\gamma}}$ has no zero in $[b, c]$. (See the algorithm below for details). Having established the initial values of $\underline{\gamma}$ and $\overline{\gamma}$, we proceed by bisection. The existence of zeros is examined for $\gamma = \frac{1}{2}(\underline{\gamma} + \overline{\gamma})$. If f_γ has zeros, then γ becomes a new value for $\underline{\gamma}$. Otherwise γ is a new value for $\overline{\gamma}$. The iteration is terminated when $\overline{\gamma} - \underline{\gamma} < \epsilon_{bis}$.

In the description of the algorithm below, we use the logical function $exist(P)$ which has the value *true* if a polynomial P has real roots in the interval $[b, c]$ and *false* otherwise.

Algorithm
Input: polynomials W and V such that $f \equiv W/V$ does have a global minimum
interval $[b, c]$
maximal error ϵ_{bis}
$x_0 \in [b, c]$
initial increment Δ^0

if $exist(W)$ **then** $\underline{\gamma} = 0$
else $\underline{\gamma} = -f(x_0)$
endif
if $\underline{\gamma} < 0$ **then** $\overline{\gamma} = 0$
else $\Delta = \Delta^0$
$\overline{\gamma} = \underline{\gamma} + \Delta$

$actex = exist(W + \overline{\gamma}V)$
until *not actex* **do**
$\Delta = 2\Delta$
$\overline{\gamma} = \overline{\gamma} + \Delta$
$actex = exist(W + \overline{\gamma}V)$
enddo
endif
{ Initial bounds are $\underline{\gamma}$ and $\overline{\gamma}$ }
until $\overline{\gamma} - \underline{\gamma} < \epsilon_{bis}$ **do**
$\gamma = (\overline{\gamma} + \underline{\gamma})/2$
if $exist(W + \gamma V)$ **then** $\underline{\gamma} = \gamma$
else $\overline{\gamma} = \gamma$
endif
enddo
{ The globally minimal value $-\gamma^\diamond$ is between the calculated values $-\overline{\gamma}$ and $-\underline{\gamma}$ }

The algorithm above is robust, and we encountered no problems in applying it even to polynomials of very high degree. However, it converges with the rate of bisection method i.e. $ln\frac{1}{\epsilon_{bis}}$. Usually this is not a problem, since about sixty steps are required to achieve the accuracy of the globally minimal value $\epsilon_{bis} = 10^{-15}$. Sometimes however, a higher accuracy is required or the degree of a polynomial is very high, and it is desirable to reduce the number of steps. Later in this section, we describe the procedure to use the much faster secant algorithm to find the globally minimal value. It works for polynomials and most rational functions but should be used with caution for some rational functions (see next section for discussion).

Having found $\overline{\gamma}$ and $\underline{\gamma}$, one can proceed to locate the global minimum position $x^\diamond$. First, we check whether the global minimum is at either of the endpoints b, c. This poses no problem because calculation of $f(b)$ and $f(c)$, and comparison with $\overline{\gamma}$ and $\underline{\gamma}$, answers the question whether the global minimum occurs at b or c.

In the following, we consider finding the position of the global minimum inside (b, c). This can be accomplished by using the Sturm chain technique for the polynomial $\underline{X}_0 \equiv W + \underline{\gamma}V$ since it has zeros at positions that differ from $x^\diamond$ only by the computational error. The algorithm is again a bisection, but this time , it divides the interval $[b, c]$ until a prescribed accuracy is achieved *i.e.* two numbers $\underline{x}^{bis-lft}$ and $\underline{x}^{bis-rgt}$ are found so that the polynomial $\underline{X}_0 = W + \underline{\gamma}V$ has zeros in $(\underline{x}^{bis-lft}, \underline{x}^{bis-rgt})$ and $\underline{x}^{bis-rgt} - \underline{x}^{bis-lft} < \epsilon_x$. In this algorithm, the Sturm chain is always the same; only values of polynomials in the chain are calculated at different points. It should be stressed that the error of determining the position of the global minimum might be larger than ϵ_x unless the number of zeros of $\underline{X}_0 = W + \underline{\gamma}V$ in (b, c) is equal to the number of zeros of this polynomial in $(\underline{x}^{bis-lft}, \underline{x}^{bis-rgt})$. If there are only two zeros of $\underline{X}_0$ in (b, c), i.e. $\underline{x}^1$ and $\underline{x}^2$, then $x^\diamond \in (\underline{x}^1, \underline{x}^2)$. The difference $| \underline{x}^1 - \underline{x}^2 |$ can be estimated from the cut Taylor expansion for f, provided that the second derivatives at $\underline{x}^{bis-lft}$ and $\underline{x}^{bis-rgt}$ are practically the same and positive. The formula used for estimation of $| \underline{x}^1 - \underline{x}^2 |$ is

$$S = \sqrt{\frac{8\epsilon_{bis}}{f''}} \tag{2}$$

where f'' is a second derivative of f at $\underline{x}^1$ or $\underline{x}^2$. We have found no problem with using the above estimate of the separation of zeros as an upper bound on the error of the position of the global minimum, even for very ill-conditioned problems. Nevertheles, one can always use the Sturm technique to locate *both* zeros of $\underline{X}_0$ to ensure correctness of error estimation.

In the following, however, we show a much faster way of locating the global minimum (minima) by examination of the last polynomial in the Sturm chain for $W + \underline{\gamma} V$. For simplicity, we start by considering the exact value $\gamma^\diamond$. The globally minimal value of the polynomial $X_0^\diamond \equiv W + \gamma^\diamond V$ is equal to 0. If there is a root $x^\diamond$ of $X_0^\diamond$ in the interior of $[b, c]$, then $x^\diamond$ is also a local minimum of $X_0^\diamond$ ($X_0^\diamond$ is nonnegative in $[b, c]$). This means that $(x - x^\diamond)$ is a common divisor of $X_0^\diamond$ and $X_1^\diamond = X_0^{\diamond\prime}$. Since the last polynomial in Sturm's chain $X_r^\diamond$ is the greatest common divisor of $X_0^\diamond$ and $X_1^\diamond$, it is divisible by $(x - x^\diamond)$. In other words, any interior root of $X_0^\diamond$ is also a root of $X_r^\diamond$. On the other hand, of course, any root of $X_r^\diamond$ is also a root of $X_0^\diamond$. Hence, all the interior global minima of f can be located by finding roots of the last polynomial $X_r^\diamond$. Practically, we do not know the exact value of $\gamma^\diamond$ hence coefficients of $\underline{X}_{r+1}$ in the Sturm chain for $\underline{X}_0 \equiv W + \underline{\gamma} V$ are small in absolute value rather than being exactly zero. Often $\underline{X}_r$ is linear; if not, its root $\underline{x}$ or roots $\underline{x}_1, \ldots, \underline{x}_s$ can be found by any standard technique for polynomials, including Sturm's chain technique. It may also be helpful to check the number of zeros of $\underline{X}_0$. If it has, for example, one zero and also $\underline{X}_0(b) < 0$, then there is no point in examining the last polynomial in the Sturm chain $\underline{X}_r$ because the global minimum is apparently at b.

The fact that the Sturm chain has fewer elements for $\gamma = \gamma^\diamond$ can be used for constructing a faster algorithm to find $\gamma^\diamond$. If $X_r^\diamond$ is the last polynomial in the Sturm chain for $\gamma = \gamma^\diamond$, then all coefficients $a_{r+1,j}(\gamma)$ of the polynomial X_{r+1} are zero for $\gamma = \gamma^\diamond$. Let $a_L(\gamma)$ be the leading coefficient of X_{r+1} calculated for $W + \gamma V$ (this is the coefficient obtained by dividing the polynomials in the Euclidean algorithm, but obviously before dividing by the absolute value of the leading coefficient). Hence, $\gamma^\diamond$ is the solution to the equation $a_L(\gamma) = 0$. This equation has many solutions [for any $\gamma = -f(x_{extr})$ where x_{extr} is any of the extremal points of the function f]. Therefore, one should start solving it from a rather good approximation of $\gamma^\diamond$. Let us assume that we start with the bisection method to initiate γ for a faster but less robust algorithm for solving $a_L(\gamma) = 0$. Bisection provides two values $\underline{\gamma}$ and $\overline{\gamma}$ such that $\underline{\gamma} < \gamma^\diamond < \overline{\gamma}$ and $\overline{\gamma} - \underline{\gamma} < \epsilon_{bis}$. If ϵ_{bis} is less than the smallest difference between different extremal values of

f, then there is usually only one zero $\gamma^\diamond$ of a_L in $(\underline{\gamma}, \overline{\gamma})$. We have chosen the secant method (Stoer and Bulirsch, 1993) to solve $a_L(\gamma) = 0$. It starts with two approximate values $\gamma_1 = \underline{\gamma}$ and $\gamma_2 = \overline{\gamma}$ and, for two subsequent approximations γ_{i-1} and γ_i, it constructs the zero point γ_{i+1} of the secant line:

$$\gamma_{i+1} = \gamma_{i-1} - a_L(\gamma_{i-1}) \frac{\gamma_i - \gamma_{i-1}}{a_L(\gamma_i) - a_L(\gamma_{i-1})}. \tag{3}$$

The convergence order of a secant method is $(1 + \sqrt{5})/2$ (Stoer and Bulirsch, 1993). We terminate the secant iteration when $\mid \gamma_i - \gamma_{i-1} \mid < \epsilon_{sec}$. Determining the correct number of polynomials in the Sturm chain for $W + \gamma^\diamond V$ is required in order to proceed with the secant algorithm. The last polynomial in the Sturm chain for $W + \gamma^\diamond V$ is generally of the form:

$$X_r^\diamond(x) = (x - x_1^\diamond)^{\sigma_1} \cdot \ldots \cdot (x - x_m^\diamond)^{\sigma_m} P(x)$$

where m is the number of global minima of f and $P(x)$ is a polynomial which does not have real zeros. For most practical cases, $P(x)$ is just a constant and all σ_i are equal to 1, but of course that should not be generally assumed. In our algorithm we used the sum of absolute values of coefficients of a polynomial in the Sturm chain as the way of determining which polynomial is going to be the last while γ tends to $\gamma^\diamond$. Having found $\underline{\gamma}$ and $\overline{\gamma}$ by the bisection algorithm, we compare the sum:

$$\eta_k(\gamma) \equiv \sum_{i=0}^{m_k} \mid \alpha_{ki}(\gamma) \mid \tag{4}$$

with ϵ_{bis}. The smallest k, for which $\eta_k(\underline{\gamma}) < \epsilon_{bis}$, is assumed to be the index of a polynomial X_{r+1} in the Sturm chain which is going to vanish for $\gamma = \gamma^\diamond$. The leading coefficient $\alpha_{r+1,m_{r+1}}$ of this polynomial is the one which is used to refine γ by solving the equation

$$a_L(\gamma) \equiv \alpha_{r+1,m_{r+1}}(\gamma) = 0$$

by the secant algorithm. However, especially for a low accuracy of the bisection initialization (large ϵ_{bis}), the correctness of the choice of r should be checked. The simple, reliable test is made by checking $\eta_{r+1}(\gamma^*)$ where γ^* is the result of secant refinement. For the correct choice of r, not only the leading coefficient of X_{r+1}^* but also all other coefficients of this polynomial should be very close to zero. (The leading coefficient $a_L(\gamma^*)$ is very close to zero by virtue of the equation being solved).

3. Numerical Examples

All examples studied were global minimizations in $(-\infty, \infty)$. We have carried out all conversions and symbolic algebra computations as well as the actual global

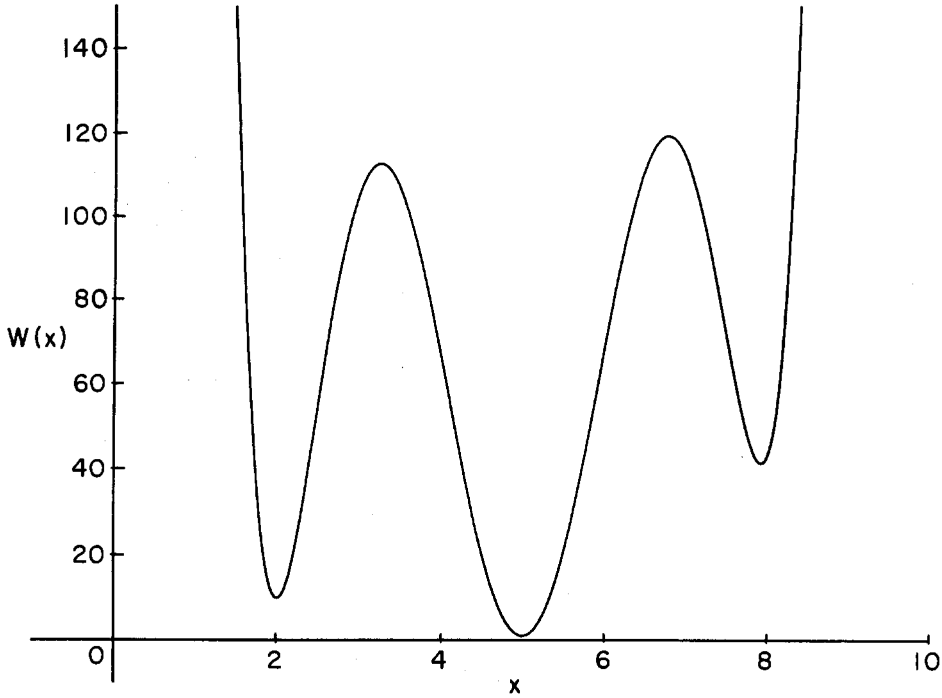


Fig. 1. Polynomial $W(x) = (x-5)^2\{(x-2)^2[(x-8)^2 + \frac{1}{10}] + 1\} + 1$.

minimization using the Maple V system, i.e. our programs were written in Maple V language. We modified the Maple V library *sturm* procedure in order to deal efficiently with infinite b and c. We also used the library *sturmseq* procedure which constructs the Sturm chain. We have encountered no problem in using Sturm chains generated directly by the Euclidean algorithm. Even for difficult examples, *sturmseq* behaved very reliably. We start by presenting a simple example, a polynomial with three different local minima:

$$\begin{aligned} W(x) &= (x-5)^2\{(x-2)^2[(x-8)^2 + \frac{1}{10}] + 1\} + 1 \qquad (5) \\ &= x^6 - 30x^5 + 357.1x^4 - 2141.4x^3 + 6763.9x^2 + \\ &\quad -10584x + 6436. \end{aligned}$$

The global minimum is obviously at $x = 5$, and the value of the polynomial at this point is 1.

The positions of all extrema and their values are listed in Table I and the function is plotted in Figure 1. We used 20-digits arithmetic. The numerical cost of the calculation will be expressed hereafter in terms of calls to the procedure generating

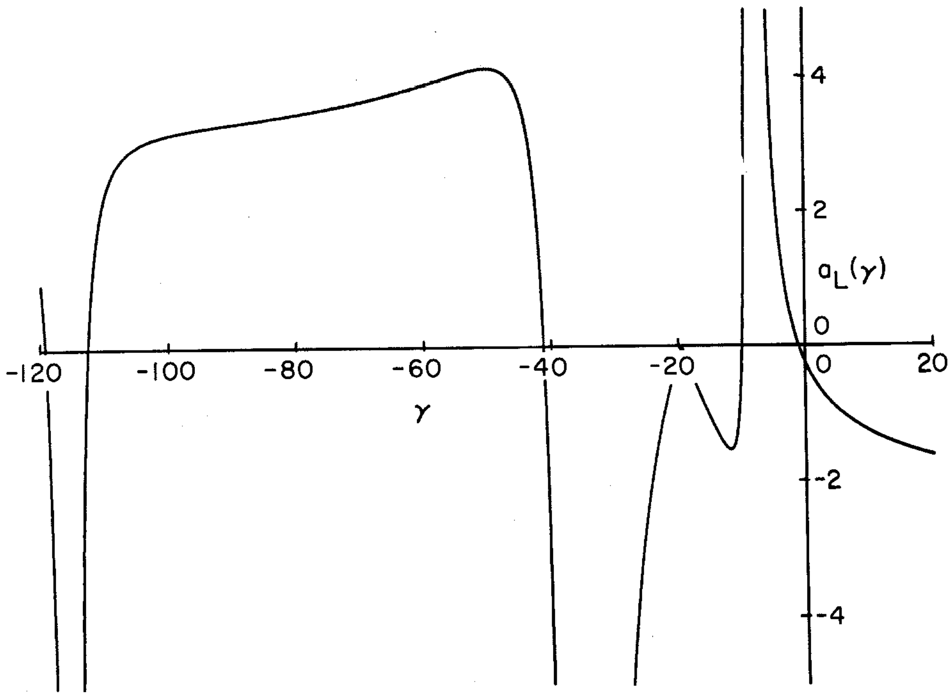


Fig. 2. The coefficient $\alpha_{6,0}$ of the Sturm chain polynomial for

$$W(x) + \gamma = (x-5)^2\{(x-2)^2[(x-8)^2 + \frac{1}{10}] + 1\} + 1 + \gamma$$

as a function of γ.

the Sturm chain named *sturmseq*. The first part of the calculation was a rough estimation of γ by the bisection algorithm. The index of the last element in the Sturm chain for $W + \gamma^{\diamond}$ was $r^{\diamond} = 5$. Then the equation $\alpha_{6,0}(\gamma) \equiv a_L(\gamma) = 0$ was solved by the secant method. The required accuracy ϵ_{bis} to terminate the bisection was set to $\epsilon_{bis} = 10^{-1}$, whereas the accuracy for the secant refinement was $\epsilon_{sec} = 10^{-15}$. The total number of calls to *sturmseq* was equal to 13 (8 in initialization and bisection and 5 in secant refinement) for $x_0 = 2$ and $\Delta^0 = 1$. The resulting globally minimal value was $-\gamma^* = 0.99999999999999976244$ ($error = 2.4 \times 10^{-16}$). The position of the global minimum as the zero point of the polynomial X_5^* in the Sturm chain for $W + \gamma^*$ was $x^* = 5.0000000000000000043$ ($error = 4.3 \times 10^{-18}$). The secant method converges fast but it often requires good initial approximations. In order to see how close the initial bisection should approach the solution of $a_L(\gamma) = 0$, we plot $a_L(\gamma) = \alpha_{6,0}(\gamma)$ in the region corresponding to all extremal values of W (Figure 2). The function is zero for all extremal values of W and also for $\gamma = -19.4958$ for which the leading coefficient $\alpha_{4,2}(\gamma)$ of the polynomial X_4 is zero. $\alpha_{6,0}(\gamma)$ is singular for γ values for which the leading coefficient $\alpha_{5,1}(\gamma)$ of

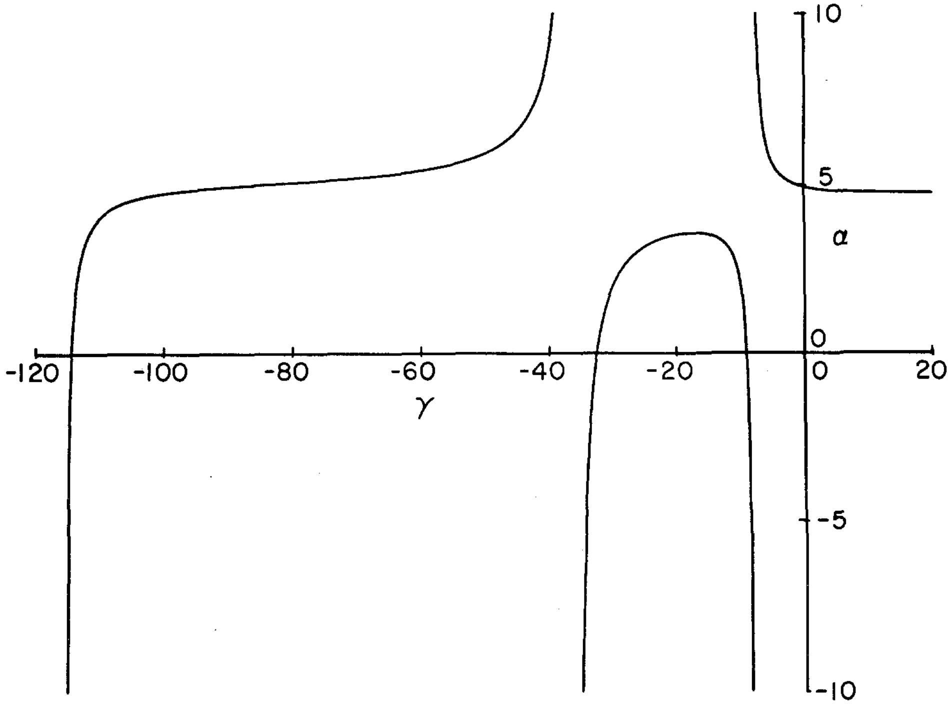


Fig. 3. The solution x of the equation $\alpha_{5,1}(\gamma)x + \alpha_{5,0}(\gamma) = 0$ as a function of γ for the Sturm chain for $W(x) + \gamma = (x-5)^2\{(x-2)^2[(x-8)^2 + \frac{1}{10}] + 1\} + 1 + \gamma$.

the polynomial $X_5 = \alpha_{5,1}x + \alpha_{5,0}$ is zero. While solving the equation $\alpha_{6,0}(\gamma) = 0$ in the vicinity of $\gamma = -1$ one should be so close to -1 so that the nearest singularity at $\gamma = -8.1678$ does not interfere. $\epsilon_{bis} = 10^{-1}$ which we used is safe, but even $\epsilon_{bis} = 1$ works fine. In order to present the sensitivity of the position of the global minimum calculated from the last polynomial in the Sturm chain on the value of γ we plot this function in Figure 3.

TABLE I. Positions and values at extrema of $W(x) = (x-5)^2\{(x-2)^2[(x-8)^2 + \frac{1}{10}] + 1\} + 1$

Position x	Value $W(x)$
2.009334882840419	9.972126485849188
3.252837828322365	112.506976264750820
5.000000000000000	1.000000000000000
6.800358695400958	119.200397429823912
7.937468593436258	41.237462782539020

The next example is a polynomial with two global minima:

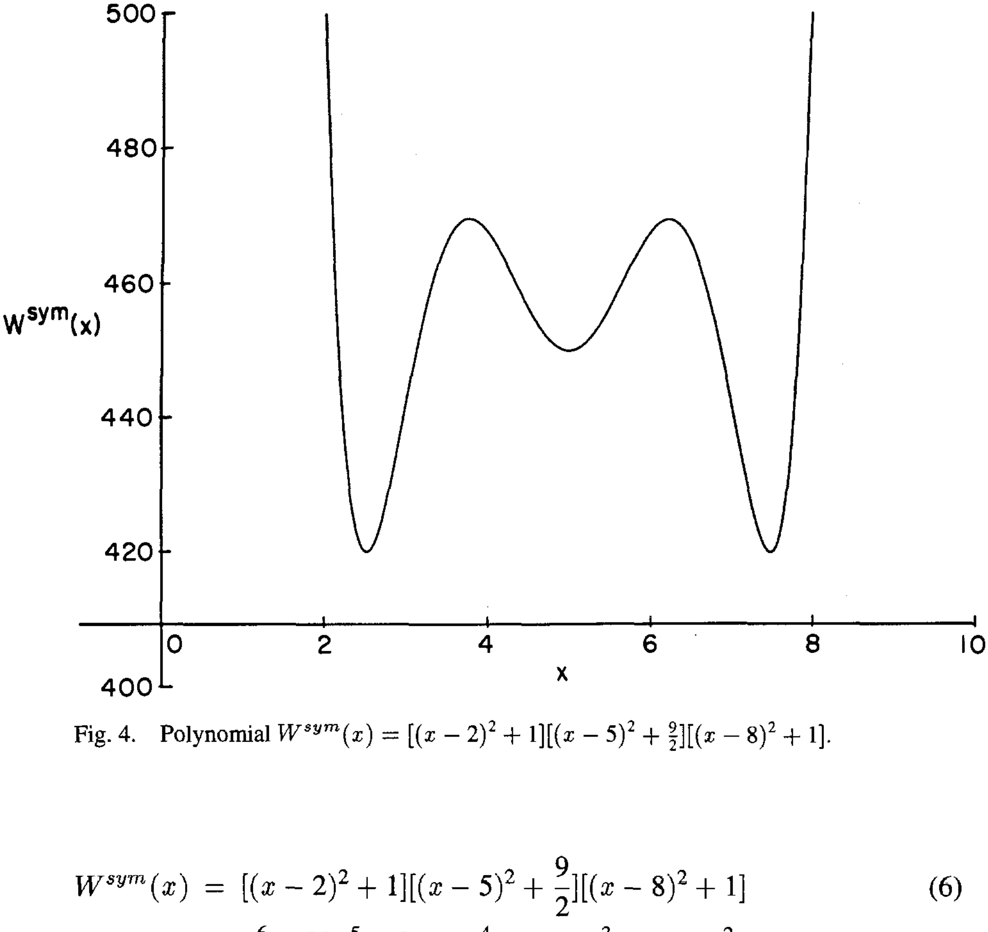


Fig. 4. Polynomial $W^{sym}(x) = [(x-2)^2+1][(x-5)^2+\frac{9}{2}][(x-8)^2+1]$.

$$\begin{aligned} W^{sym}(x) &= [(x-2)^2+1][(x-5)^2+\frac{9}{2}][(x-8)^2+1] \\ &= x^6 - 30x^5 + 363.5x^4 - 2270x^3 + 7678x^2 + \\ & -13280x + 9587.5 \end{aligned} \tag{6}$$

The positions of all extrema and their values are presented in Table II. A plot of the polynomial is presented in Figure 4.

Again, 20-digit arithmetic was used, the accuracies of the bisection and the secant were $\epsilon_{bis} = 10^{-1}$ and $\epsilon_{sec} = 10^{-15}$, respectively, and $x_0 = 5$ and $\Delta^0 = 1$. The index of the last polynomial in the Sturm chain for $\gamma = \gamma^\diamond$ for this example was assumed to be $r^\diamond = 4$. After 19 calls to *sturmseq* (14 in initialization and bisection and 5 in secant) the globally minimal value $-\gamma^* = 419.8496139818166$ was obtained. The positions of the two global minima had errors of 2.57×10^{-16} and 2.98×10^{-16}. The polynomial X_5^* in the Sturm chain for $W^{sym} + \gamma^*$ was equal to $X_5^* = -1.36 \times 10^{-16}x - 3.96 \times 10^{-15}$. It, obviously, had a very small leading coefficient because that was the criterion to find γ^*, but also the second coefficient was very small which was a numerical confirmation that $r^\diamond$ was chosen

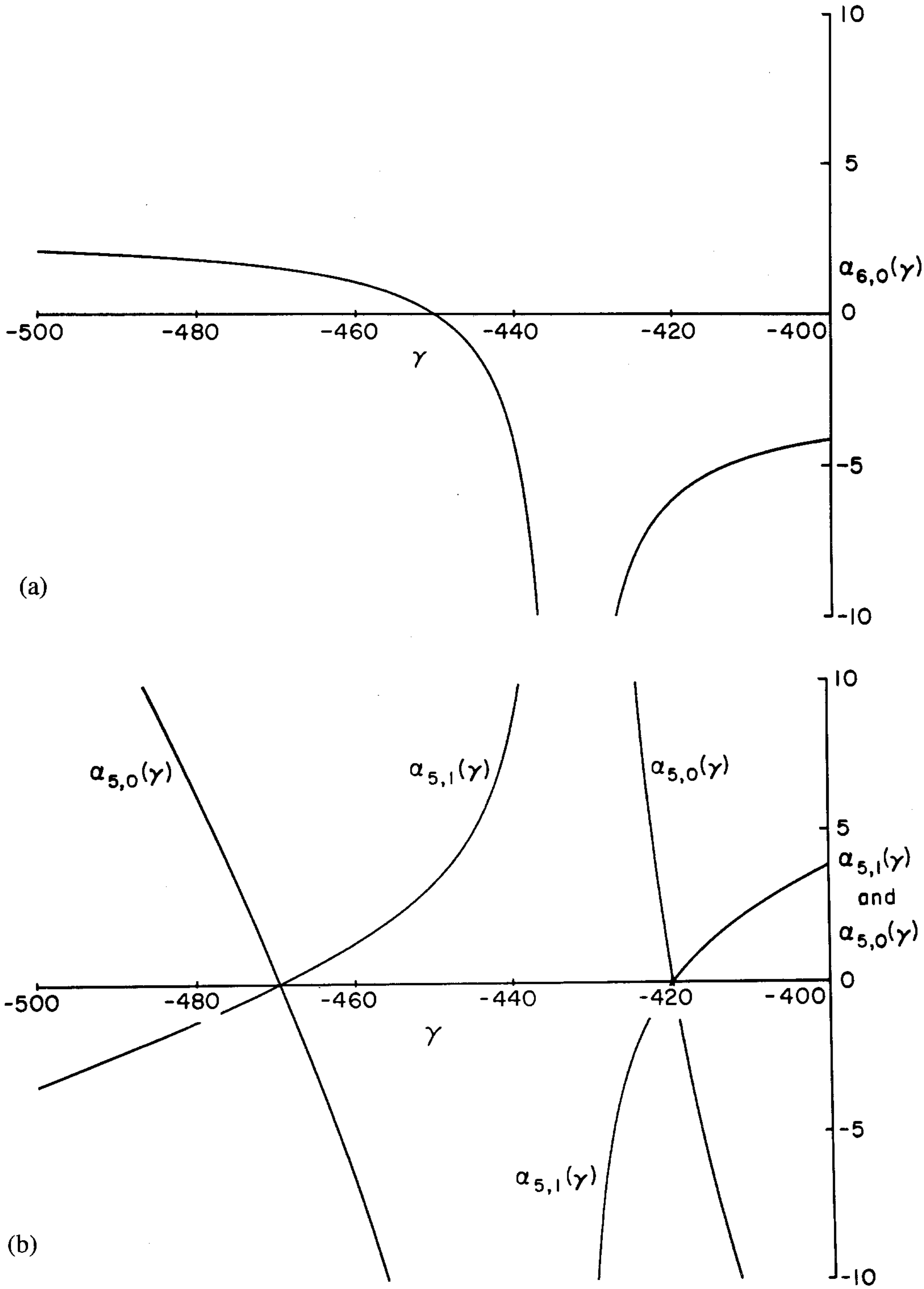


Fig. 5. Coefficients of the two last polynomials in the Sturm chain for

$$W^{sym}(x) + \gamma = [(x-2)^2 + 1][(x-5)^2 + \frac{9}{2}][(x-8)^2 + 1] + \gamma$$

(a) $\alpha_{6,0}(\gamma)$; and (b) $\alpha_{5,1}(\gamma)$ and $\alpha_{5,0}(\gamma)$.

TABLE II. Positions and values at extrema of $W^{sym}(x) = [(x-2)^2+1][(x-5)^2+\frac{9}{2}][(x-8)^2+1]$

Position x	Value $W^{sym}(x)$
2.520334565932673	419.8496139818166
3.767958603877951	469.5022378700353
5.000000000000000	450.0000000000000
6.232041396122048	469.5022378700353
7.479665434067326	419.8496139818166

correctly. (The fact that this polynomial has two global minima implies that degree of $X_r^\diamond$ should be at least two and it was equal to 2 but generally it could be greater than 2). We have studied the dependence of the coefficient $\alpha_{5,0}^*$ on the actual difference between the values at the lowest and the second lowest minimum of the polynomial. In Table III, the polynomial X_5^* is presented for several polynomials $W_\zeta \equiv W^{sym} + \zeta x$ minimized in the same way as W^{sym} was minimized previously. It is clear that the value of $\alpha_{5,0}^*$ is a good test for the separation of the two lowest minima. In the same table we show errors for positions of the two lowest minima of W_ζ calculated as zeros of X_4^*. In Figure 5 we show plots of the functions $\alpha_{6,0}(\gamma)$, $\alpha_{5,1}(\gamma)$ and $\alpha_{5,0}(\gamma)$. The only zero of $\alpha_{6,0}(\gamma)$ corresponds to the single central minimum of W^{sym}. Both $\alpha_{5,1}(\gamma)$ and $\alpha_{5,0}(\gamma)$ have zero value for γ corresponding to the double global minimum and the double maximum of W^{sym}.

TABLE III. Influence of the difference between minimal values of W_ζ assumed to be identical, on the polynomial X_5^* and on the errors of positions of these minima calculated as zeros of X_4^*

Value of ζ	Distance between the two lowest minimal W_ζ values	X_5^* for W_ζ	Error of x_1^*	Error of x_2^*
10^{-9}	4.96×10^{-9}	$7.4 \times 10^{-17}x - 2.08 \times 10^{-9}$	1.22×10^{-10}	1.22×10^{-10}
10^{-5}	4.96×10^{-5}	$9.0 \times 10^{-17}x - 2.08 \times 10^{-5}$	1.22×10^{-6}	1.22×10^{-6}
10^{-1}	4.96×10^{-1}	$1.6 \times 10^{-16}x - 0.208$	1.21×10^{-2}	1.23×10^{-2}

The next group of examples demonstrates how the method performs for high-degree rational functions. We have tested a group of functions which are of practical importance, truncated Fourier series. Before global minimization by our algorithm, the Fourier series needs to be converted into a rational function. This can be achieved by expanding sines and cosines of multiple angles in terms of sines and cosines of single angles and, then, by substitution:

$$\sin\alpha = \frac{2x}{1+x^2}$$

$$\cos\alpha = \frac{1-x^2}{1+x^2}$$

where $x = \tan\frac{\alpha}{2}$; this establishes a mapping from $\alpha \in (-\pi, \pi)$ into $x \in (-\infty, \infty)$. Obviously, the result of this substitution is a rational function which has a global minimum in $\mathcal{R}$. We used four different truncated Fourier series of the form

$$F(\alpha) = \sum_{n=1}^{N}(a_n \cos n\alpha + b_n \sin n\alpha) \tag{7}$$

The number of terms was chosen as $N = 18$, and the Fourier coefficients were generated according to the following equations:

Example F1

$$a_n = \sin(2n+9)$$

$$b_n = \sin(7n-3)$$

Example F2

$$a_n = \sin(2n^2+6)$$

$$b_n = \sin(6n+4)$$

Example F3

$$a_n = \sin(2n+6)$$

$$b_n = \sin(4n-3)$$

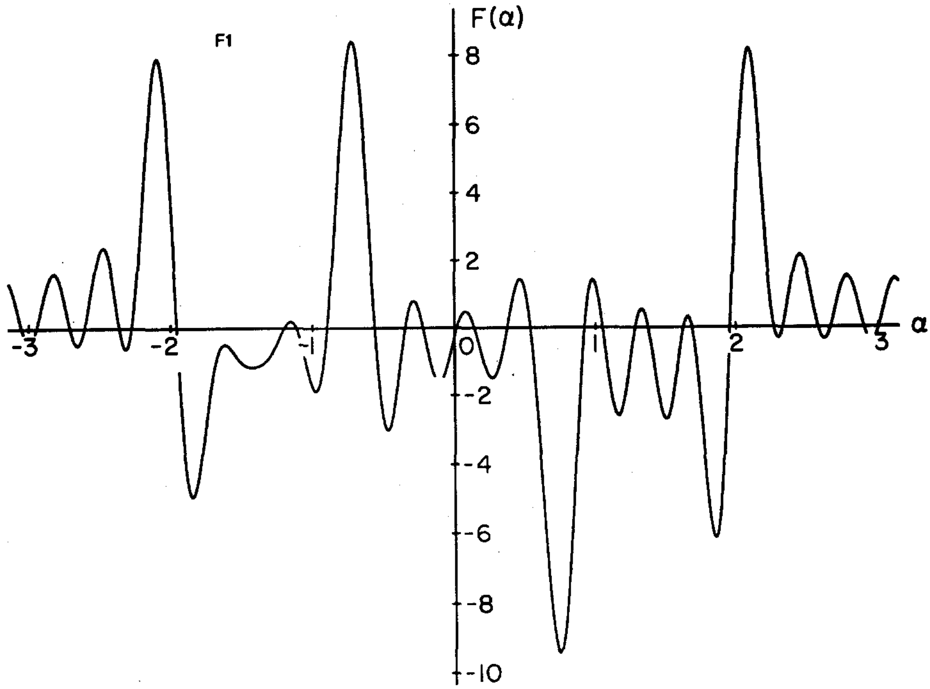
Example F4

$$a_n = \sin(3^n)$$

$$b_n = \sin(7^n).$$

We do not list here the actual polynomials W and V for these functions since they are defined uniquely by the Fourier coefficients. The polynomials are all of degree 36.

Plots of function (7) for these four choices of Fourier coefficients are shown in Figure 6, and the globally minimal values obtained with 20-digit arithmetic and the accuracy of determining the globally minimal value $\epsilon_{bis} = 10^{-15}$ are collected in Table IV. The number of calls to *sturmseq* procedure was 57 for each example. We used the robust bisection algorithm to obtain reliable minimal values. The results for the secant method are presented later in this section.



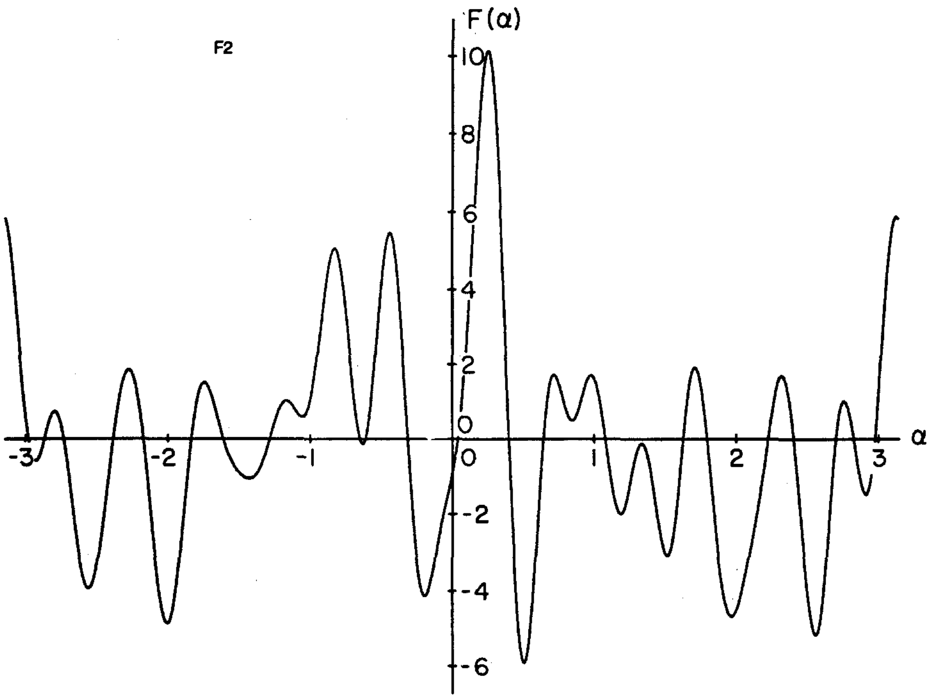


Fig. 6. Truncated Fourier series F1,F2,F3 and F4.

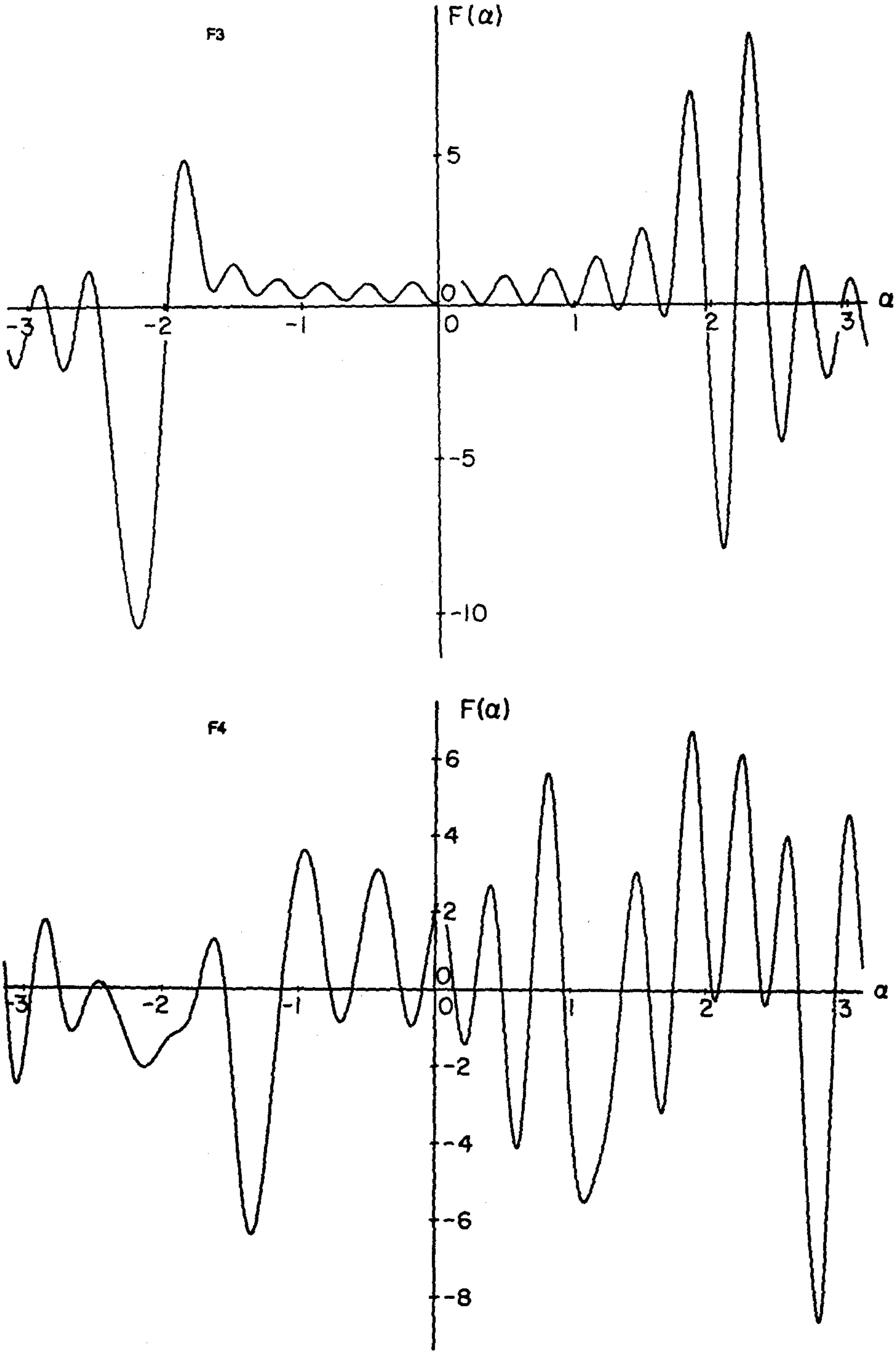


Fig. 6. Continued.

TABLE IV. Globally minimal values, for $\epsilon_{bis} = 10^{-15}$, 20-digits arithmetic, $N = 18$, $x_0 = 0$ and $\Delta^0 = 4$. For every example, the number of calls to *sturmseq* was equal to 57

Example	Globally minimal value
F1	-9.427079657662615
F2	-5.904604685207378
F3	-10.447409929498691
F4	-8.587921046512747

TABLE V. Comparison of the global minimum position calculated from the last polynomial in the Sturm chain and by the bisection algorithm

Example	$\underline{\alpha} =$	Bisection	
	$= 2arctan(\underline{x})$	Calls to sturm	$\underline{\alpha}^{bis} = 2arctan(\underline{x}^{bis-lft})$
F1	0.7367992888	53	0.7367992888
F2	0.4977600271	53	0.4977600262
F3	0.0003575575	55	-2.1996890318
F4	-0.0115422839	59	2.8169550460

The position of the global minimum was calculated both from the last element of the Sturm chain ($\underline{\alpha}$) and also by the bisection algorithm described in section 2 ($\underline{\alpha}^{bis}$). The cost of the bisection algorithm is expressed in terms of the number of calls to the procedure *sturm* which calculates values of polynomials in the Sturm chain at the endpoints of an interval. In using two methods, we aimed to establish how safe it is to rely on the last element of the Sturm chain for locating the global minimum of a rational function. See Table V for results. The estimated separation S of zeros of $f_{\underline{\gamma}}$ from formula (2) is presented in Table VI. Since ϵ_x in the bisection algorithm

TABLE VI. Separation of zeros of $f_{\underline{\gamma}}$ and the value of the denominator polynomial V at zero of $W + \underline{\gamma}V$

Example	Separation of zeros of $f_{\underline{\gamma}}$, S	$V(\underline{x}^{bis-lft})$
F1	2.96×10^{-9}	1.22×10^{1}
F2	2.68×10^{-9}	3.09×10^{0}
F3	4.36×10^{-9}	2.27×10^{12}
F4	2.41×10^{-9}	3.13×10^{28}

was chosen as 10^{-15}, S is a measure of the error of $\underline{\alpha}^{bis}$. For examples F1 and F2, the root of the last (linear) polynomial in the Sturm chain $\underline{X}_{35}$ is in good agreement with the position of a zero of the polynomial $\underline{X}_0 = W + \underline{\gamma} V$ found by the bisection algorithm. However, for the other two examples, the accuracy $\epsilon_{bis} = 10^{-15}$ for determining the globally minimal value is insufficient to establish the position of the global minimum from $\underline{X}_{35}$. The reason for this is that the denominator polynomial V has a very large value at the zero point $\underline{x}$ of the polynomial $\underline{X}_0 = W + \underline{\gamma} V$ [see Table VI for the values of $V(\underline{x})$]. The algorithm terminates when $\overline{\gamma} - \underline{\gamma} < \epsilon_{bis}$ so that the minimal value of $f_{\underline{\gamma}}$ cannot differ from 0 by more than ϵ_{bis}, but the value of the polynomial $W + \underline{\gamma} V = f_{\underline{\gamma}} V$ at $\underline{x}$ may be very large when $V(\underline{x})$ is large. When $\epsilon_{bis} = 10^{-15}$ in examples F3 and F4, the value of $W + \underline{\gamma} V$ at $\underline{x}$ may be as large as 2×10^{-3} and 3×10^{13} respectively. Hence, it is necessarry to set ϵ_{bis} to a very small value in order to be able to learn the position of the global minimum from $\underline{X}_r$. Table VII illustrates this by showing positions for the examples F3 and F4 for higher accuracies. $\epsilon_{bis} = 10^{-35}$ is sufficient to obtain the correct position for example F3 but not for example F4. Only to prove that the behavior for example F4 is due to insufficient accuracy ϵ_{bis}, we repeated the calculations for $\epsilon_{bis} = 10^{-70}$ and obtained the correct position of the global minimum from $\underline{X}_r$. Since increasing accuracy for determining the globally minimal value is expensive (see number of calls to the *sturmseq* procedure in Table VII), we believe that the cheaper and safer way of locating the global minimum for ill-conditioned problems such as F3 and F4 is to use the bisection algorithm [one call to *sturm* is usually faster than a call to *sturmseq*, although checking for the existence of zeros in $(-\infty, \infty)$ requires only checking the sign of leading coefficients without evaluating any polynomial].

TABLE VII. Positions of the calculated global minima from $\underline{X}_r$ for different values of ϵ_{bis}

Example	Number of digits	ϵ_{bis}	Calls to sturmseq	$\underline{\alpha} = 2arctan(\underline{x})$	$\underline{\alpha}^{bis} = 2arctan(\underline{x}^{bis-lft})$
F3	40	10^{-35}	123	-2.1996890318	-2.1996890318
F4	40	10^{-35}	123	-0.0115422839	2.8169550466
F4	80	10^{-70}	240	2.8169550466	2.8169550466

The number of calls to the *sturmseq* procedure depends on the initial difference between the bounds $\overline{\gamma} - \underline{\gamma}$ and the required accuracy ϵ_{bis} for finding the globally minimal value. Since every bisection step divides $\overline{\gamma} - \underline{\gamma}$ by a factor of 2, the number of steps n, necessary to achieve the accuracy ϵ_{bis}, is the smallest integer n satisfying the inequality:

$$2^{-n}(\overline{\gamma} - \underline{\gamma}) < \epsilon_{bis}$$

or equivalently:

$$n > \frac{1}{\ln 2} \ln \frac{\overline{\gamma} - \underline{\gamma}}{\epsilon_{bis}}.$$

In other words, the cost of the method depends on ϵ_{bis} as $\ln \frac{1}{\epsilon_{bis}}$.

We also tested the secant algorithm for the rational functions F1, .. ,F4. We learned that, for the examples for which there is no problem in locating the global minimum from the last Sturm chain polynomial, there is also no problem in finding the globally minimal value by the secant method. For the example F1, using the secant algorithm reduced the number of calls to *sturmseq* from 57 to 22 and, for the example F2, the number of calls was reduced from 57 to 25. However, for examples F3 and F4, the required initial accuracy ϵ_{bis} to start the secant method and obtain the globally minimal value was at the order of the required accuracy to obtain a good position of the global minimum from the last element in the Sturm chain. Hence, for these examples, the bisection with $\epsilon_x = 10^{-15}$ was faster.

The computational cost for each call of *sturmseq* should be proportional to the square of the degree of the polynomial $W + \gamma V$. The reason for this is that the number of terms in the Sturm chain scales linearly with the degree of the polynomial and that the number of coefficients to be calculated for every polynomial in the Sturm chain is equal to its degree. To confirm the above estimation, we carried out calculations for different numbers of terms N in the Fourier series for example F2. (See Table VIII for the timings). Similarly, the *sturm* procedure scales as the square of the degree of a polynomial because the cost of the calculation of the value of a polynomial scales linearly with its degree.

TABLE VIII. Execution time on an IBM RISC 6000 computer as a function of the number of terms for example F2. Execution time involves finding the initial bounds $\overline{\gamma}$ and $\underline{\gamma}$ as well as bisection for γ.

Number of terms	Calls to sturmseq	Execution time [sec]
6	54	22
12	57	74
18	57	151
36	59	643

Acknowledgements

We thank Dr. R. Wawak for helpful remarks about this manuscript. This work was supported by a grant from the National Science Foundation (DMB90-15815). Support was also received from the Association for International Cancer Research.

References

Cetin, B. C., Barhen, J. and J. W. Burdick (1993). Terminal Repeller Unconstrained Subenergy Tunnelling (TRUST) for Fast Global Optimization. *Journal of Optimization Theory and Applications* **77**: 97-126.

Hansen, E. R. (1979). Global Optimization Using Interval Analysis: The One-Dimensional Case. *Journal of Optimization Theory and Applications* **29**: 331-344.

Stoer J. and R. Bulirsch (1993). *Introduction to Numerical Analysis*. Springer-Verlag: New York, p. 308.

van der Waerden, B. L. (1991). *Algebra*. Springer-Verlag: New York, Vol. I, pp. 244-245.

Visweswaran, V. and C. A. Floudas (1992). Unconstrained and Constrained Global Optimization of Polynomial Functions in One Variable. *Journal of Global Optimization* **2**: 73-99.

Wingo, D. R. (1985). Globally Minimizing Polynomials without Evaluating Derivatives. *International Journal of Computer Mathematics* **17**: 287-294.