APPROXIMATION OF COMPLEX FUNCTIONS BY VECTOR PROJECTION
USING LEAST MEAN-SQUARE METHODS

W. Göpfert
Siemens AG, Geschäftsbereich Weitverkehrstechnik
D-8000 München, FR. Germany

Abstract

A novel method for the approximation of complex functions even in the case of relatively poor starting values of the approximation-parameters for using least mean-square methods is described. The method relies largely on the introduction of an appropriate approach by vector projection for reducing complex variables to real variables.

Introduction

In computer-aided circuit design in general and in designing microwave filters in particular the approximation of given curves by functions with suitably determinable parameters is assuming steadily increasing importance. For adapting real functions to series of scalar variables there are numerous well-tried methods[1-5] for minimizing the sum of the squares of the deviations between the ideal values and the values assigned to the approximating function.

It is also becoming increasingly necessary to use complex functions for approximating series of complex values of for instance the reflection coefficient or transmission coefficient of a filter. If good starting values of the function parameters to be determined are already available, this problem can, owing to the quasi linear response of the complex network functions to small changes in the parameters, be solved by minimizing the sum of the magnitudes of the differences between the given and the approximating complex values. Thereby the well-tried least mean-square methods can also be used for approximating complex functions. If however only relatively poor parameter starting values are available for the approximation, these methods will often fail in the sense of yielding a secondary minimum of the objective function. Such failure is clearly due to the choice of the magnitude of the difference between complex values as the variable to be minimized if the complex approximating function has no longer a quasi linear response.

1. Approximation Of Complex Functions By Minimizing The Mean Square Amplitude Error

For approximating a given series of complex values

$$\underline{r}^g_l = r^g_{Rl} + j\, r^g_{Il} \quad , \quad l = 1,2,\ldots,m$$

by the values

$$\underline{r}^a_l = \underline{r}^a_l(p_i) = r^a_{Rl} + j\, r^a_{Il} \quad ,$$
$$l = 1,2,\ldots,m;\; i = 1,2,\ldots,n$$

of a complex function $\underline{r}^a(p_i)$, dependent on the parameters p_i, those parameters p_i^* are to be determined, which satisfy the two requirements

$$r^a_{Rl} - r^g_{Rl} \overset{!}{=} 0 \quad \text{and} \quad r^a_{Il} - r^g_{Il} \overset{!}{=} 0$$
$$\text{or} \qquad\qquad l=1,2,\ldots,m \qquad (1)$$
$$\underline{r}^a_l - \underline{r}^g_l \overset{!}{=} 0 \;.$$

We see, that in consequence of the relations

$$|\underline{r}^1 - \underline{r}^2| = \sqrt{(r^1_R - r^2_R)^2 + (r^1_I - r^2_I)^2}$$

and

$$\sqrt{(r^1_R - r^2_R)^2 + (r^1_I - r^2_I)^2} = 0 \;\Leftrightarrow\; r^1_R - r^2_R = 0 \,\cap\, r^1_I - r^2_I = 0$$

the two requirements (1) can be substituted by the equivalent requirement

$$D_l = |\underline{r}^a_l - \underline{r}^g_l| \overset{!}{=} 0 \qquad (2)$$

Thereby the problem of approximating a series of complex values can be reduced to the approximation of a series of real values, and it can be manipulated by using some of the well-known methods for minimizing the mean-square error, whereby the minimization problem

$$\sum_{l=1}^{m} |\underline{r}^a_l(p_i) - \underline{r}^g_l|^2 \overset{!}{=} \text{MIN} \qquad (3)$$

has to be solved.

Since complex network functions show an quasi linear response to minor variations in parameters[6], the variation of the parameters p_i by $\Delta p_i \ll p_i$ will result in the terminal

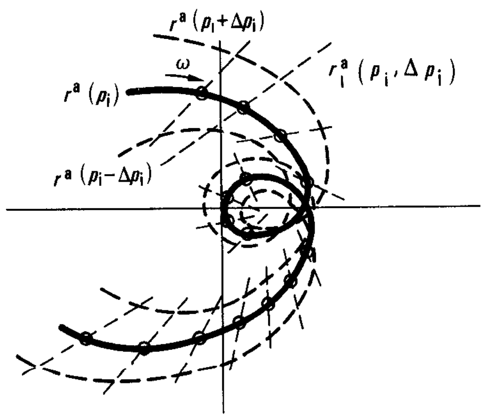

Fig. 1: Quasi linear response of the reflection coefficient to minor variations $\Delta p_i \ll p_i$.

points of the vectors wandering along the geometric locus approximately represented by straight lines (Fig. 1), and by appropriately varying the parameters it is possible to bring the vectors $\underline{r}_l^a$ and $\underline{r}_l^g$ into congruence by decremental reduction of the intervalls $|\underline{r}_l^a - \underline{r}_l^g|$. For larger values of Δp_i this procedure is not longer possible on account of the increasingly nonlinear response.

2. Definition Of A Real Approximation For The Complex Approximation Error By Vector Projection

If the series of complex values are represented as series of vectors in the complex plane, the problem of approximation will consist in bringing the respective vectors of the approximation function $\underline{r}^a(p_i)$ into congruence with the given series $\underline{r}_l^g$ by varying the parameters p_i (Fig. 2).

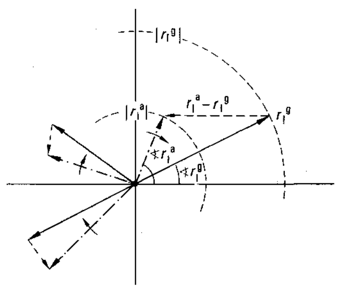

Fig. 2: Vectorial representation of a given series of complex values $\underline{r}_l^g$ and their approximation by $\underline{r}_l^a$.

For congruence the two requirements

$$\Delta\sphericalangle_l = \sphericalangle\underline{r}_l^a - \sphericalangle\underline{r}_l^g \overset{!}{=} 0 \qquad l=1,2,..,m \qquad (4)$$
$$|\Delta r|_l = |\underline{r}_l^a| - |\underline{r}_l^g| \overset{!}{=} 0$$

must be satisfied. Fig. 2 shows the equivalence of the requirements (4) and (2).

Varying the parameters by Δp_i usually causes the rotation of each vector $\underline{r}_l^a$ by $\Delta\sphericalangle\underline{r}_l^a$ accompanied by a variation in magnitude of $\Delta|\underline{r}_l^a|$. For an approximation the parameters have to be varied such, that the vectors can be rotated to the desired value $\sphericalangle\underline{r}_l^g$ and the magnitude can be approximated to $|\underline{r}_l^g|$.

As an appropriate combination of the two requirements (4) into a single requirement relating to one objective variable, the following approximation is introduced

$$D_l^T = |\underline{r}_l^a|\cdot|\cos\Delta\sphericalangle_l|^k\cdot\mathrm{sign}(\cos\Delta\sphericalangle_l) - |\underline{r}_l^g| \overset{!}{=} 0 \qquad (5)$$

This approximation can be interpreted as follows: Project the approximating vector $\underline{r}^a$ by the regulation $|\underline{r}_l^a|\cdot|\cos\Delta\sphericalangle_l|^k\cdot\mathrm{sign}(\cos\Delta\sphericalangle_l)$ in a directional line segment s_l to the straight line g passing through $\underline{r}_l^g$. For a rotation of $\underline{r}_l^a$ the line segment s_l on the straight line g assumes positive and negative values of a magnitude which depends on both $|\underline{r}_l^a|$ and $\sphericalangle\underline{r}_l^a$.

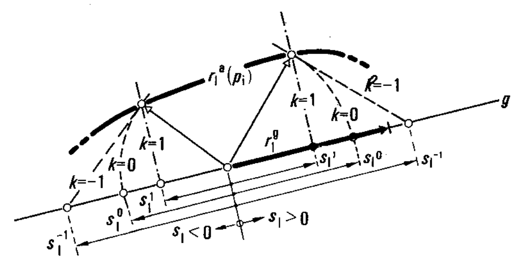

Fig. 3: Projection s_l with various selected values of k.

Fig. 3 gives the resulting values of s_l for various chosen values of k and especially for k=0. In this case the requirement $D_l^T \overset{!}{=} 0$ (5) leads to the expression

$$|\underline{r}_l^a|\cdot\mathrm{sign}(\cos\Delta\sphericalangle_l) - |\underline{r}_l^g| \overset{!}{=} 0 ,$$

which essentially contains the requirement for the congruence between the magnitudes of the vectors, additionally however includes the condition $-\pi/2 < \Delta\sphericalangle_l < \pi/2$ by taking into consideration the angle between the vectors $\underline{r}_l^a$ and $\underline{r}_l^g$.

For all k>0 values the variable s_l passes through a value range of an upper bound $s_{max}^- < 0$ to a lower bound $s_{max}^+ > 0$ whereby for all $\underline{r}_l^a$ with $\sphericalangle\underline{r}_l^a = \pi/2, 3\pi/2$ independent of $|\underline{r}_l^a|$ the zero is passed through and the values s_{max}^- and s_{max}^+ are yielded by the resulting $|\underline{r}_l^a|$. For all values of k<0 the variable s_l passes through a value range of $s_{min}^+ > 0$ to $+\infty$ and $-\infty$ to $s_{min}^- < 0$, whereby for $\sphericalangle\underline{r}_l^a = \pi/2, 3\pi/2$ independent of $|\underline{r}_l^a| > 0$ the value of $+\infty$ changes to $-\infty$ or vice versa and the values s_{min}^- and s_{min}^+ are likewise yielded by the resulting $|\underline{r}_l^a|$.

It is obvious that the requirement (5) does not automatically yield $|\Delta\underline{r}_l^a|=0$ and $\Delta\sphericalangle_l=0$, and consequently $\underline{r}_l^a = \underline{r}_l^g$. Fig. 4 shows some geometric loci of permissible "optimum vectors" $\underline{r}_{l\,opt}^a$ for various values of k. The appearance of such geometric loci of permissible vectors is a general result of the combination of the two requirements (4) into one approximation requirement (5).

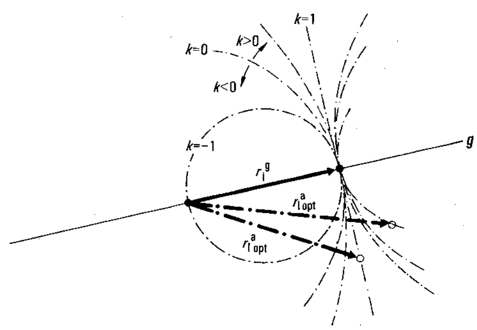

Fig. 4: Geometric loci of permissible "optimum vectors" $\underline{r}_{l\,opt}^a$ as a function of k.

3. Approximation Of Complex Functions By Minimizing The Sum Of The Squares Of D_l^T

Using the newly introduced approximation $D_l^T \overset{!}{=} 0$ istead of the two requirements $|\underline{r}_l^a| \overset{!}{=} 0$ and $\Delta\sphericalangle_l=0$, the problem of the approximation of a series of vectors $\underline{r}_l^g$ by the complex function $\underline{r}^a(p_i)$ is reduced to the minimization problem

$$\sum_{l=1}^{m}(s_l - |\underline{r}_l^g|)^2 \overset{!}{=} \text{MIN}$$

and consequently to the approximation of real values.

The approximation process can graphically be interpreted as follows: Using the projection s_l, which depends on both amplitude and angle of the vectors, a rotation of the approximating vectors $\underline{r}_l^a$ into the required vectors $\underline{r}_l^g$ is effected.

Fig. 3 shows, that all values of $k<0$ are only efficient, if the condition $|\Delta\measuredangle_l|<\pi/2$ holds for all $\underline{r}_l^a$. For all $\underline{r}_l^a$ with $|\Delta\measuredangle_l|\geq\pi/2$ the reduction of the angle $\Delta\measuredangle_l$ between $\underline{r}_l^a$ and $\underline{r}_l^g$, that means a correction of the approximation, namely leads to an increasing value of s_l and would announce the change for the worse in the approximation process. The required assumption $|\Delta\measuredangle_l|<\pi/2$, $l=1,2,..,m$ however cannot be taken generally. Moreover the values $k<0$ are unsuitable on account of the discontinuities of s_l at $\pi/2$ and $3\pi/2$.

The influence of either the magnitude or the angle of the vectors on the approximation process can be increased with the aid of k. In particular it is possible to define an exponent $k^*= k(\Delta\measuredangle)$ dependent on $\Delta\measuredangle$ which results in the angle becoming more pronounced for large deviations $\Delta\measuredangle$ and the magnitude becoming more pronounced for small $\Delta\measuredangle$. The expression

$$k^* = (2 - \cos\Delta\measuredangle_l)\cdot k_o$$

has here proved very effective.

Owing to the permissible geometric loci of optimum vectors it is also essentially possible to find a minimum MIN=0 for the case $\underline{r}_l^a \neq \underline{r}_l^g$. This solution is however known from experience to come so close to the solution $\underline{r}_l^a = \underline{r}_l^g$, that the response of $\underline{r}^a(p_i)$ is already linear, whence the minimization of the mean square error of the amplitudes of the vector differences (3) is achieved in only a few steps.

The proposed method has been successfully used in conjunction with a least mean-square method described by D.W.Marquard[5] for the approximation of reflection coefficients and transmission coefficients, especially in cases where the minimization of the amplitudes of the differences between the vectors $\underline{r}_l^a$ and $\underline{r}_l^g$ led to remote secondary minima. The number of optimizing steps was in all cases significantly reduced, whereby k=1 was generally used.

Fig. 5 shows with reference to an example successive approximation by minimizing the mean square error of the amplitudes of the vector differences (3). The best approximation obtained as the final result of four optimizing steps is still far remote from the desired optimum and obviously represents a local minimum.

For the same example Fig. 6 shows the improvements achieved in six successive optimizing steps through the use of the new method. The result of the last step corresponds with the ideal curve.

Conclusion

By the introduction of the proposed approximation for the complex approximation error it is now possible to solve the problem of approximation of complex functions using the well-tried least mean-square methods even in the case of relatively poor starting values of the approximation parameters. The advances as to the state-of-art are an increased region of solvability and a substantial reduction of the numerical effort.

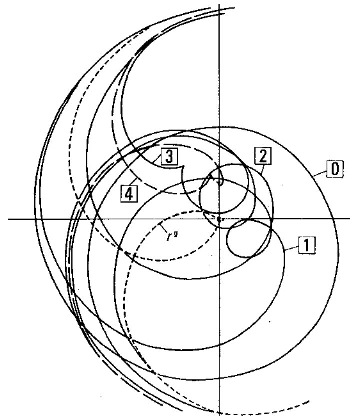

Fig. 5: Approximation of $\underline{r}_l^g$ using $\sum_{l=1}^{m}|r_l^a-r_l^g|^2 \overset{!}{=} \text{MIN}$ Starting values [0] and four optimizing steps [1]-[4]. Curve [4], the final result of the optimizing process, is still far remote from the ideal curve. (secondary minimum!)

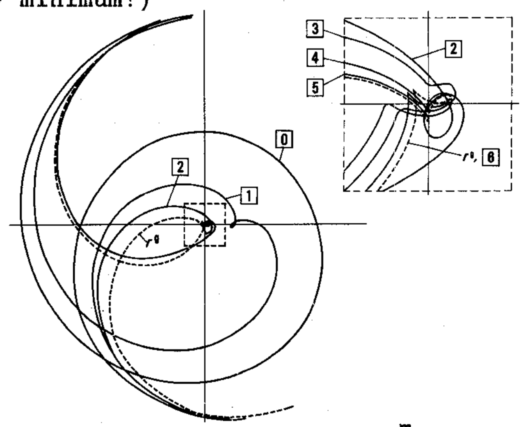

Fig. 6: Approximation of $\underline{r}_l^g$ using $\sum_{l=1}^{m}(s_l-|r_l^g|)^{2} \overset{!}{=} \text{MIN}$ Starting values [0] and six optimizing steps [1]-[6]. Curve [6], the result of the last optimizing step, corresponds with the ideal curve $\underline{r}_l^g$.

References

1 C.G.Broyden, "Quasi-Newton methods and their application to function minimization". Math. Comp. 21 (1967), pp. 368-381

2 R.Fletcher, M.J.D.Powell, "A rapidly convergent descent method for minimization". Comp. J. 6 (1963), pp. 163-168

3 K.Levenberg, "A method for the solution of certain nonlinear problems in least squares". Quart. Appl. Math. 2 (1944), pp. 164-168

4 M.R.Osborn, "Some aspects of non-linear least-square calculations". In: F.A.Lootsma Ed., Numerical methods for non-linear optimization, Academic Press, London/New York 1972

5 D.W.Marquard, "An algorithm for least squares estimation of non-linear parameters". SIAM 11 (1963), pp. 431-441

6 H.Gutsche, "Fast tolerance analysis of electrical network using multidimensional random variables". IEEE Intern. Symp. CAS Proc. (1976), pp. 718-721