Q-factor measurement with a scalar network analyser

D. Kajfez

Indexing terms: Electrical properties, Filter resonators, Coupling coefficients, Reflection coefficients, Q-factor

Abstract: An accurate determination of the unloaded Q-factor of RF and microwave resonators is required when the electrical properties of materials are being measured or when the resonator is to be characterised for an application in a filter or an oscillator. A procedure is described of weighted least-square curve fitting to the measured magnitude of the reflection coefficient. The accuracy of the results is significantly improved by incorporating coupling losses in the equivalent circuit model.

1 Introduction

The need for an accurate measurement of the unloaded Q-factor arises in the following situations:

(*a*) determination of the material properties of a dielectric or conductor sample inside a resonant cavity

(*b*) determination of the circuit properties of a resonator that is to be incorporated into a filter or an oscillator.

When the measurement is performed with an automatic network analyser, the unloaded Q and the coupling coefficient are most accurately obtained by the data-fitting procedures [1–3]. For the data measured by a vector network analyser, the linear fractional data fitting [4, 5] provides a convenient and accurate approach to graphical display and data processing operations. In some situations, only the scalar network analyser is available, providing the magnitude, but not the phase, of the reflection coefficient. It has been shown [3] that it is possible to determine the unloaded Q-factor even in that situation. However, the results become inaccurate when coupling losses are present in the experimental setup. This paper describes an improved data processing procedure which is appropriate for moderate coupling losses.

2 Equivalent circuit

The magnitude of the input reflection coefficient Γ is measured at a number of frequencies in the vicinity of the resonant frequency. The input reflection coefficient can be expressed as the following function of frequency ω [4, eqn. 220]:

$$\Gamma = \Gamma_d + \frac{de^{j\gamma}}{1 + jQ_L 2\dfrac{\omega - \omega_L}{\omega_0}} \tag{1}$$

The resonant frequency of the unloaded resonator is denoted by ω_0 and that of the loaded resonator is denoted by ω_L. For all practical purposes, ω_0 in the denominator can be replaced by ω_L. At the outset of the data processing procedure, ω_L is found by determining the frequency at which $|\Gamma|$ is minimum. The loaded Q-factor is denoted Q_L, the diameter of the Q-circle is denoted d, and the detuned input reflection coefficient is denoted Γ_d (a complex number):

$$\Gamma_d = \rho e^{j\varphi} \tag{2}$$

An equivalent circuit that corresponds to eqn. 1 is shown in Fig. 1. The unloaded resonator is represented by a

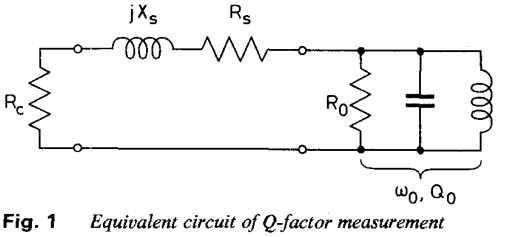

Fig. 1 *Equivalent circuit of Q-factor measurement*

parallel resonant circuit, characterised by the unloaded Q-factor Q_0, unloaded resonant frequency ω_0, and resistant R_0 which stands for the power dissipated within the unloaded resonator. The external loading circuit ($\equiv$network analyser) is characterised by the resistance R_c. The coupling mechanism which connects the external circuit to the unloaded resonator is characterised by an impedance $R_s + jX_s$. The resistance R_s represents the losses within the coupling mechanism, and the reactance X_s represents the extra energy storage introduced by the coupling mechanism.

For a lossy coupling mechanism $\rho < 1$, so that the Q-circle does not touch the circumference of the Smith chart, as indicated in Fig. 2. The squared magnitude of

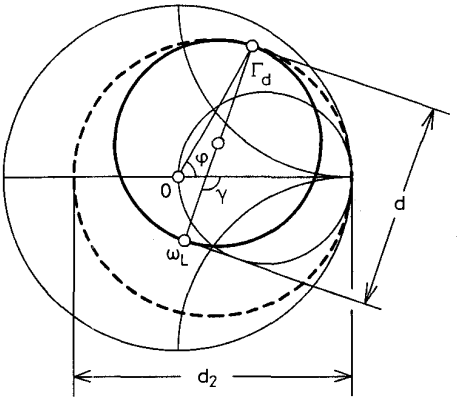

Fig. 2 *Input reflection coefficient against frequency*

Paper 2142H (E12), received 24th March 1995 and in final revised form

The author is with the University of Mississipi, Department of Electrical Engineering, Anderson Hall, University, MS 38677, USA

the input reflection coefficient in eqn. 1 can be expressed as follows:

$$-a_1|\Gamma|^2t^2 + a_2t^2 + a_3 + a_4t = |\Gamma|^2 \tag{3}$$

where the coefficients a_1 and a_4 are

$$a_1 = Q_L^2 \tag{4}$$

$$a_2 = Q_L^2\rho^2 \tag{5}$$

$$a_3 = \rho^2 + 2\rho d\cos\theta + d^2 \tag{6}$$

$$a_4 = -2\rho\, dQ_L \sin\theta \tag{7}$$

The relative frequency has been denoted by t

$$t = 2\frac{\omega - \omega_L}{\omega_0} \tag{8}$$

It is seen from eqns. 4–7 that the as are expressed in terms of circuit parameters to be measured: Q_L, d, ρ, and θ. In principle, when the data of N measured frequencies are substituted in eqn. 3, one obtains a linear system of N equations, to be solved for the four coefficients a_1 *to* a_4. This is done by the least square data fitting procedure. Once the values of the four coefficients are determined, the quantities of principal interest are easily computed.

The angle θ is obtained in the process of evaluating the squared magnitude of eqn. 1 as follows

$$\theta = \varphi - \gamma \tag{9}$$

It can be seen in Fig. 2 that φ is the phase of Γ_d, and γ is the phase of the vector of the length d, pointing from the tip of Γ_d through the center of the Q-circle. The difference of these two angles θ is an angle slightly smaller than π. As long as the coupling losses are small,

$$\cos\theta \simeq -1 \quad \text{and} \quad \sin\theta \simeq 0 \tag{10}$$

Therefore, $a_4 \simeq 0$, so the system can be treated as consisting of only three unknown coefficients. Also, a_3 simplifies to

$$a_3 \simeq (\rho - d)^2 \tag{11}$$

3 Weighted least-square solution

For each of the measured frequencies, the values $|\Gamma|$ and t may be substituted into eqn. 3. An overdetermined system of linear equations is created that is solved in much the same way as recommended in classical literature on the subject [6–8]. The procedure is here modified by using weighted scalar products as in [4, 5], so that each measurement is assigned a different weight. For the scalar measurements considered here, this is done in the following way.

When the data of the ith measurement are substituted in eqn. 3, the result is not exactly zero, but is equal to some small number, ε_i

$$-a_1|\Gamma_i|^2t_i^2 + a_2t_i^2 + a_3 + a_4t_i - |\Gamma_i|^2 = \varepsilon_i \tag{12}$$

Some measured points will result in a larger error, others in a smaller one. A statistically proper procedure is to assign to each measurement a weight inversely proportional to the squared error. If it is assumed for a moment that a_1 to a_4 are variables, and $\sigma^2(a_1)$ to $\sigma^2(a_4)$ are their variances, the error propagation formula [9] gives the variance of ε_i as follows

$$\sigma^2(\varepsilon_i) = |\Gamma_i|^4t_i^4\sigma^2(a_1) + t_i^4\sigma^2(a_2) + \sigma^2(a_3) + t_i^2\sigma^2(a_4) \tag{13}$$

Therefore the weight to be assigned to the ith measurement is

$$p_i = \frac{\sigma^2}{|\Gamma_i|^4t_i^4\sigma^2(a_1) + t_i^4\sigma^2(a_2) + \sigma^2(a_3) + t_i^2\sigma^2(a_4)} \tag{14}$$

The constant σ^2 in the numerator is any convenient number, the value of which cancels out in the final result. A convenient choice is $\sigma^2 = \sigma^2(a_3)$ so that at the resonant frequency ($t = 0$) the weight is equal to unity. At the outset of computation, the variances $\sigma^2(a_1)$ to $\sigma^2(a_4)$ are not known. The solution of the system then proceeds in several iterations. In the first iteration, the weights are evaluated by setting all variances equal to unity. The solution of the system provides the values of all four coefficients and their variances. These are then used in the second iteration to compute the weights, etc. The process converges very rapidly: third and further iterations do not show a noticeable effect on the final results of the data fitting.

The solution of the overdetermined system (eqn. 3) has been performed by two different approaches: solving the full system consisting of four unknown coefficients, and solving the approximate system with the simplifying assumption of eqn. 10 that results in only three unknown coefficients. It has been found that in most cases both approaches give very similar final results, but the four-term approach is more prone to numerical difficulties. In the simpler, three-term approach, the quantities of interest are evaluated as follows.

The loaded Q is computed from eqn. 4:

$$Q_L = \sqrt{(a_1)} \tag{15}$$

The magnitude of the detuned reflection coefficient is obtained from eqns. 4 and 5:

$$\rho = \sqrt{\left(\frac{a_2}{a_1}\right)} \tag{16}$$

There are two possible values of the Q-circle diameter, obtained from eqn. 11:

$$d = \rho \pm \sqrt{(a_3)} \tag{17}$$

Such an ambiguity does not occur in the vector formulation of the problem [4, 5], because in that case, the linear fractional curve fitting procedure results in a single-valued solution for d.

The coupling losses are specified by the diameter d_2 of the 'coupling loss circle' indicated in Fig. 2 by a dashed line. In that Figure, the losses are exaggerated for better display. In the case of small coupling losses, diameter d_2 is only one or two percent shorter than the diameter of the Smith chart. Then d_2 can be approximated as

$$d_2 \simeq 1 + \rho \tag{18}$$

The coupling coefficient κ that takes into account the coupling losses is then computed by [4, eqn. 229]:

$$\kappa = \frac{1}{\frac{d_2}{d} - 1} \tag{19}$$

Finally, the unloaded Q-factor is

$$Q_0 = Q_L(1 + \kappa) \tag{20}$$

4 Measured results

Before applying the procedure to an actual set of measured data, a number of simulated data sets has been

generated for an assumed equivalent circuit from Fig. 1. After the procedure was developed to the point that it was capable of accurately recovering the assumed values of Q_0 and κ, it was applied to the measured data. The example presented consists of the measured input reflection coefficient of an aluminum cavity that contains a dielectric resonator [10]. The actual measurement was performed by the vector network analyser. The data file, named M450.MEA, has been processed by the linear fractional data fitting procedure, and the results are ([14], p. 130):

$$Q_L = 4536.9 \quad Q_0 = 8083.0 \quad \kappa = 0.7816$$

For the purpose of demonstrating the curve fitting procedure for the scalar measurements, only the magnitude of the measured reflection coefficient is retained and the phase ignored. The three-term approach gives the results shown in Fig. 3. It is seen that the overcoupled $Q_0 =$

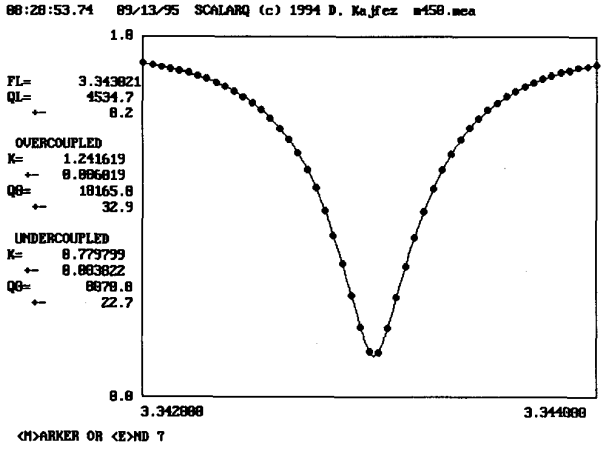

Fig. 3 *Results of three-term curve fitting*

8070.8, agreeing within 0.15% with the value obtained from the vector network analyser data.

Fig. 3 is a 'print-screen' reproduction of the monitor display obtained by the program called SCALARQ*. The dark circles are the measured data. In this measurement, there are 51 data points, centred around the frequency 3.34 GHz. The solid line is the fitted curve, computed from the coefficients a_1 to a_3. Each of the estimated parameters printed to the left of the plot also contains an estimated error that has been computed from variances $\sigma^2(a_1)$ to $\sigma^2(a_3)$.

It can be seen that the measured data allow two possible interpretations: one undercoupled ($\kappa < 1$) and the other overcoupled ($\kappa > 1$). From the measured data alone, it is not possible to conclude which of the two interpretations is the correct one. In the present example the results of another measurement are known (namely those obtained by the vector network analyser), so one can tell that the correct interpretation is the undercoupled one. If the scalar network analyser is the only available instrument, some additional experimental verification is needed to decide whether one is dealing with an overcoupled or an undercoupled case. For instance, the coupling probe may be pulled slightly out of the cavity and the entire measurement repeated. One can expect two possible outcomes: the undercoupled κ will either increase or decrease. In the first case, one may conclude that the proper interpretation is the overcoupled one, in the last the undercoupled one.

* The DOS version of the program SCALARQ accompanies the book '*Q* factor' [4]

It is interesting to observe the values of the individual weights for this example, plotted in Fig. 4. They have been computed by eqn. 14 after three iterations of the

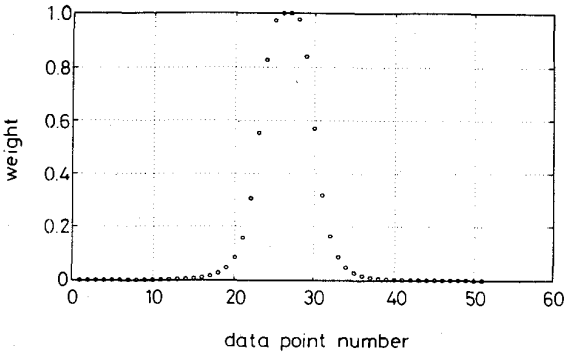

Fig. 4 *Set of weights for example from Fig. 3*

curve fitting procedure. The weights are normalised, so the largest weight is unity. Clearly, the weights strongly favour the measured data closest to the resonant frequency of the loaded resonator.

5 Comparison with two-term method

The curve fitting procedure described in [3] can also process the data obtained by the scalar network analyser. The procedure does not use weighting and the equivalent circuit is assumed to have a lossless coupling. Therefore the magnitude of the detuned reflection coefficient is assumed to be $\rho = 1$, so that $a_4 = 0$, $a_3 = (1 - d)^2$, $a_1 = a_2 = Q_L^2$. Only two distinct coefficients remain to be fitted, namely a_1 and a_3. When such a two-term procedure is applied to the same set of measured data (file M450.MEA), the least-square fitted curve does not agree well with the measured data, as can be seen in Fig. 5. The

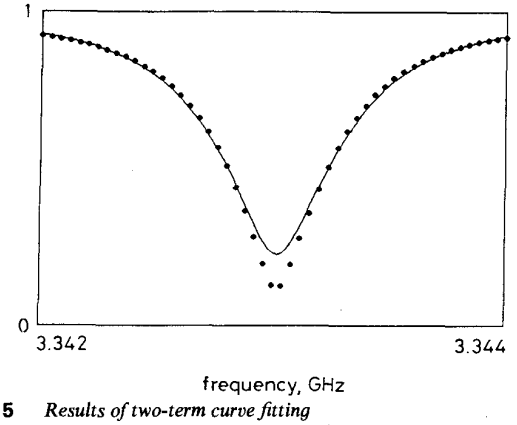

Fig. 5 *Results of two-term curve fitting*

unloaded Q-factor computed by the two-term method is $Q_0 = 6663$, a value which is 17% lower from the one obtained by the three-term method. This example demonstrates that an overly simplified equivalent circuit may lead to a significant loss of accuracy.

6 Conclusions

For an accurate determination of the unloaded Q-factor with the scalar network analyser, the magnitude of the input reflection coefficient is measured at a number of frequencies. When the measured data are substituted into a circuit model of the resonator with coupling losses an overdetermined system of equations with four unknown coefficients is obtained. For moderate coupling losses the system can be reduced to three unknown coefficients.

This paper describes a three-term procedure of determining the coupling coefficient and the unloaded Q-factor. As a side benefit, the procedure also provides the error estimates for the measured quantities. If the coupling losses are ignored, the system reduces to a two-term procedure, and the accuracy of the results deteriorates.

An inherent problem of measuring the Q-factor with a scalar network analyser is the fact that there are two possible interpretations for any set of measured data, one is an overcoupled circuit and the other is an undercoupled one. A separate test must be performed to determine which of the two interpretations is correct. The test proposed in this paper is to slightly decrease the physical coupling and then repeat the measurement.

7 References

1 WHELESS, W. P., and KAJFEZ, D.: 'Microwave resonator circuit model from measured data fitting'. 1986 IEEE MTT-S symposium digest, Baltimore, June 1986, pp. 681–684

2 SANCHEZ, M.C., MARTIN, E., and ZAMARRO, J.: 'New vectorial automatic technique for characterisation of resonators', *IEEE Proc. H.*, 1989, **136,** (2), pp. 145–150

3 SANCHEZ, M.C., MARTIN, E., and ZAMARRO, J.M.: 'Unified and simplified treatment of techniques for characterising transmission, reflection or absorption resonators', *IEE Proc. H.*, 1990, **137,** (4), pp. 209–212

4 KAJFEZ, D.: 'Q-factor' (Vector Fields, 1994)

5 KAJFEZ, D.: 'Linear fractional curve fitting for measurement of high Q-factors', *IEEE Trans.*, 1994, **MTT-42,** pp. 1149–1153

6 WHITTACKER, E.T., and ROBINSON, G.: 'The calculus of observations' (Blackie, 1937)

7 LINNIK, YU.V.: 'Method of least squares and principles of the theory of observations' (Pergamon, 1961)

8 GUEST, P.G.: 'Numerical methods of curve fitting' (Cambridge University Press, 1961)

9 PARRAT, L.G.: 'Probability and experimental errors in science' (Wiley, 1961)

10 GUNDAVAJHALA, A.: 'Material properties of shielded dielectric resonators by combined network analyzer measurement and finite integration technique'. PhD dissertation, Department of Electrical Engineering, University of Mississippi, August 1993, Univ. Microfilm 9406646