Musings on Circuit Q

Andrzej G. Lozowski[1], Peter B. Aronhime[2], and Tongfeng Qian[2]
1) ECE Department, Southern Illinois University, Edwardsville, IL 62026, USA
2) ECE Department, University of Louisville, Louisville, KY 40292, USA

***Abstract*—Most second-order continuous-time signal processing circuits, whether voltage-mode, current-mode, or mixed-mode, should be capable of realizing Qs greater than 1/2. However, a number of otherwise interesting circuits have been proposed in the literature which do not have this attribute. In this paper, we show a simple procedure for determining whether Q can be made greater than 1/2 for certain circuits. For other circuits, we develop a more involved procedure for this determination.**

I. Introduction

The locations of the roots of a second-order real polynomial can be described in rectangular form by the pair (α, β) or in a polar form by the pair (Q, ω_0). Thus, using Q and ω_0, the polynomial takes the form

$$K(s^2 + s\frac{\omega_0}{Q} + \omega_0^2) \tag{1}$$

where K is a scale factor which has no effect on pole locations and will be dropped in the following discussion. The constant ω_0 is the undamped natural frequency (considering the polynomial to be a denominator of a rational network function). Of course, it is well known that if $0 < Q < 1/2$, the roots of (1) are negative and unequal; and if $1/2 < Q < \infty$, the roots are conjugate complex in the open left-half plane [1]. In filter applications and other signal processing uses, conjugate pairs of poles are most often needed for transfer functions. That is, the denominators of useful transfer functions are usually built up as products of second-order polynomials having conjugate complex roots. Thus, Qs are most often needed in such applications that are greater than 1/2. However, every so often, circuits are proposed for filter applications which cannot realize complex poles, or, if they can realize complex poles, can only realize such poles that are barely complex.

In the following, we propose two methods to examine expressions for Q. The first is a simple algebraic procedure that works for some expressions but not all. The second method is more complicated, but it is more general. Examples of the applications of the methods are provided.

II. Simple algebraic method

This method is based on the fact that Q is a dimensionless quantity. First, express Q as a function of ratios of circuit elements of the same type. For example, suppose the denominator of a second-order transfer function is given by

$$s^2 + s\left(\frac{G_5}{C_5} + \frac{(C_4+C_5)(G_3G_6 - G_2G_7)}{C_4C_5G_6}\right) + \frac{G_5(G_3G_6 - G_2G_7)}{C_4C_5G_6} \tag{2}$$

where Gs are circuit conductances, and Cs are capacitances. Since the circuit should be stable, we must have $G_3G_6 > G_2G_7$. After some algebraic manipulations, we obtain

$$Q = \frac{1}{\sqrt{\frac{C_4}{C_5}}\frac{\sqrt{G_5G_6}}{\sqrt{G_3G_6 - G_2G_7}} + \left(\sqrt{\frac{C_4}{C_5}} + \sqrt{\frac{C_5}{C_4}}\right)\frac{\sqrt{G_3G_6 - G_2G_7}}{\sqrt{G_5G_6}}} \tag{3}$$

where all the ratios in the denominator of Q are dimensionless. Next, we simplify the equation for Q. For example, in (3) we let

$$x = \sqrt{\frac{C_4}{C_5}} \qquad y = \frac{\sqrt{G_3G_6 - G_2G_7}}{\sqrt{G_5G_6}} \tag{4}$$

The next step is to search for an isolated function F in the denominator of Q having the form $F = f + 1/f$ where F is a real-valued function of the circuit elements. By isolated function, we mean that F is not multiplied by any other function of circuit components. In the example, we have

$$Q = \frac{1}{\frac{x}{y} + \frac{y}{x} + xy} \tag{5}$$

where $F = x/y + y/x$ and both x and y are positive. Since the minimum of F is 2 and xy is positive, the maximum value of Q that can be realized with this circuit is less than 1/2. Thus, if Q can be expressed as 1 over a denominator, the isolated function F can be found in the denominator as illustrated in the example, and the denominator consists of F plus other positive terms, then Q must be less than 1/2. Of course, if F is multiplied by a function of components with values greater than one, then the maximum of Q will be even smaller than 1/2. On the other hand if F is multiplied by a function with values less then one or if the other functions in the denominator have negative values, it is possible that Q may be greater than 1/2.

III. Lagrange multipliers method

In this method a general search for the maximum of Q in multiple dimensions is employed. Suppose the expression for Q is given by

$$Q = \frac{\sqrt{2G_0G_2C_2(2C_1 + C_3)}}{C_2G_0 + 2G_2(C_1 + C_2 + C_3)} \tag{6}$$

This expression for Q is not amenable to the algebraic procedure given in the previous section. Variables C_1, C_2, C_3, G_0, and G_2 are real-valued and bounded in the range $(0, \infty)$. Definitely, function Q attains a maximum either inside or on the specified boundaries. Let the variables be renamed first: $x_1 = C_1$, $x_2 = C_2$, $x_3 = C_3$, $x_4 = G_0$, and $x_5 = 2G_2$. Vector $x = (x_1, x_2, x_3, x_4, x_5)$ is an element of R^5 and function $Q : R^5 \to R$ in terms of x_i's reads

$$Q = \frac{\sqrt{x_2x_4x_5(2x_1 + x_3)}}{x_2x_4 + x_5(x_1 + x_2 + x_3)} \tag{7}$$

Note that all the variables x_i are positive and so is Q in the region considered where the maximum of Q is sought. For the sake of analysis to follow, a variable change will considerably ease the task. Let $v : R^5 \to R^5$ be the following vector function:

$$v = v(x) : \begin{bmatrix} e^{v_1} \\ e^{v_2} \\ e^{v_3} \\ e^{v_4} \\ e^{v_5} \end{bmatrix} = \begin{bmatrix} 2x_1 + x_3 \\ x_2 \\ x_2x_4 + x_5(x_1 + x_2 + x_3) \\ x_4 \\ x_5 \end{bmatrix} \tag{8}$$

Even though this formulation does not provide an explicit definition of function $v(x)$, it is sufficient for the purpose of finding maxima of $Q[v(x)]$. Important properties of variable change (8) will become evident after inspecting its Jacobian. Examine the first row of the set of equations (8): Differential $\mathrm{d}e^{v_1} = e^{v_1}\mathrm{d}v_1$ from the left side of the equation is equal to the differential $\mathrm{d}e^{v_1} = \frac{\partial}{\partial x_1}(2x_1 + x_3)\mathrm{d}x_1 + \frac{\partial}{\partial x_3}(2x_1 + x_3)\mathrm{d}x_3 = 2\mathrm{d}x_1 + \mathrm{d}x_3$ as derived from the right side of the equation. This establishes the relationship between differentials in the original and transformed space $\mathrm{d}v_1 = e^{-v_1}(2\mathrm{d}x_1 + \mathrm{d}x_3)$. Finally, using this result, the first row of the Jacobian matrix can be identified from the equation $\mathrm{d}v_1 = 2\mathrm{d}x_1/(2x_1 + x_3) + \mathrm{d}x_3/(2x_1 + x_3)$. By following the same procedure for all the other equations contained in (8) and setting $\alpha = 2x_1 + x_3$ and $\beta = x_2x_4 + x_5(x_1 + x_2 + x_3)$, the Jacobian of function $v(x)$ is found to be

$$\frac{\partial v}{\partial x} = \begin{bmatrix} \frac{2}{\alpha} & 0 & \frac{1}{\alpha} & 0 & 0 \\ 0 & \frac{1}{x_2} & 0 & 0 & 0 \\ \frac{x_5}{\beta} & \frac{x_4+x_5}{\beta} & \frac{x_5}{\beta} & \frac{x_2}{\beta} & \frac{x_1+x_2+x_3}{\beta} \\ 0 & 0 & 0 & \frac{1}{x_4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{x_5} \end{bmatrix} \tag{9}$$

The Jacobian matrix resulting from these calculations is quite sparse and finding its determinant is a fairly easy task

$$\left|\frac{\partial v}{\partial x}\right| = \frac{1}{\alpha\beta x_2x_4} \tag{10}$$

The determinant (10) never attains zero in the region of positive x_i's. Therefore, the variable transformation $v(x)$ is a one-to-one mapping. Moreover, the inverse transformation $x(v)$ exists everywhere in the region considered.

The necessary condition for function $Q(x)$ to attain an extremum at a point x^* is that its gradient $\frac{\partial Q}{\partial x}$ due to coordinates x is zero at x^*. Considering the introduced variable transformation, Q can be written in terms of variables v as $Q(x) = Q[v(x)]$. Using the chain rule, the relationship between the gradient in the old and the new coordinates is

$$\frac{\partial Q}{\partial x} = \frac{\partial v}{\partial x}^T \frac{\partial Q}{\partial v} \tag{11}$$

Note that this relationship is a set of linear equations. Since the Jacobian matrix (9) is full-rank in the region of interest, the necessary condition for extrema looks similar in both coordinate systems

$$\left|\frac{\partial v}{\partial x}\right| \neq 0 \quad \Rightarrow \quad \frac{\partial Q}{\partial x} = \vec{0} \quad \Longleftrightarrow \quad \frac{\partial Q}{\partial v} = \vec{0} \tag{12}$$

Following this conclusion, the search for maximum can be performed for function $Q(v)$ and then the localized maximum v^* can be found in the original coordinates as $x(v^*)$. The variable change (8) was designed to simplify the function $Q(v)$

$$Q(v) = e^{\left(\frac{v_1}{2}+\frac{v_2}{2}-v_3+\frac{v_4}{2}+\frac{v_5}{2}\right)} \tag{13}$$

The gradient of Q due to v is proportional to $Q(v)$ itself and equals $\frac{\partial Q}{\partial v} = Q(v)(1/2, 1/2, -1, 1/2, 1/2)$. Clearly, function Q does not vanish, $Q(v) > 0$, which implies that the gradient does not vanish either $\frac{\partial Q}{\partial v} \neq \vec{0}$. Therefore, according to equivalence (12), function $Q(x)$ does not attain a maximum inside the specified region of positive x_i's.

Each x_i represents a physical parameter such as conductance or capacitance. The initial assumption was that these parameters are bounded within $(0, \infty)$. There are certain physical limitations on the set of values each of the parameters can take. Let X_{iL} and X_{iU} be the lower and the upper bound of parameter x_i, such that $(X_{iL}, X_{iU}) \subset (0, \infty)$. The overall search for the maximum of function $Q(x)$ should be performed with constraints $X_{iL} < x_i < X_{iU}$. Since the function $Q(x)$ does not attain a maximum inside boundaries, the maximum must be found on them. Therefore, the task becomes the search with constraint $x_i = X_i$, where the constant X_i represents either the lower or upper bound for x_i.

Define five constraint functions $g_i : R^5 \to R$ as follows:

$$g_i(x) = x_i - X_i \tag{14}$$

A constraint is satisfied if the corresponding constraint function equals zero. Each of the constraints determines a submanifold of R^5 on which to find a maximum [2]. Images of these submanifolds in the transformation $v(x)$ represent the constraints in the transformed coordinates v_i. In order to find those, the inverse transformation $x(v)$ is needed first:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} e^{v_1} + e^{v_2} + e^{v_2+v_4-v_5} + e^{v_3-v_5} \\ e^{v_2} \\ -e^{v_1} - 2e^{v_2} - 2e^{v_2+v_4-v_5} + 2e^{v_3-v_5} \\ e^{v_4} \\ e^{v_5} \end{bmatrix} \tag{15}$$

With the help of the inverse transformation, the constrain functions (14) in coordinates v_i can be written as:

$$\begin{aligned} g_1(v) &= e^{v_1} + e^{v_2} + e^{v_2+v_4-v_5} + e^{v_3-v_5} - X_1 \\ g_2(v) &= e^{v_2} - X_2 \\ g_3(v) &= -e^{v_1} - 2e^{v_2} - 2e^{v_2+v_4-v_5} + 2e^{v_3-v_5} - X_3 \\ g_4(v) &= e^{v_4} - X_4 \\ g_5(v) &= e^{v_5} - X_5 \end{aligned} \tag{16}$$

In the transformed coordinates, each constraint is a submanifold $g_i^{-1}(0) \subset R^5$ of codimension 1. The intersection thereof in R^5 is a point. The intersection of m of them is a submanifold of codimension m. Thus, a criterion as to which of the constraints are important from the point of view of searching for the maximum of function Q must be established.

Assume, that function $Q(v)$ attains a maximum at a point v^* on a submanifold $g_i^{-1}(0)$. This implies the level surface (isocline) $Q(v) = \text{const}$ being tangent to the submanifold at that point. In other words, Q has zero growth in the vicinity of point v^* on the submanifold. This, in sequel, implies that the gradient of Q is orthogonal to the submanifold at v^*; thus the gradient of $Q(v)$ and the gradient of $g_i(v)$ are linearly dependent at point v^*.

Similarly, assume that function $Q(v)$ attains a maximum at a point on the intersection of two submanifolds: $v^* \in g_i^{-1}(0) \cap g_j^{-1}(0)$. In this case, gradient $\frac{\partial}{\partial v} Q(v^*)$ is a linear combination of gradients $\frac{\partial}{\partial v} g_i(v^*)$ and $\frac{\partial}{\partial v} g_j(v^*)$. In general, if function $Q(v)$ attains a maximum at v^* on a submanifold of codimension m, then the gradient of Q is a linear combination

$$\frac{\partial Q(v^*)}{\partial v} = \sum_{i=1}^{5} \lambda_i^* \frac{\partial g_i(v^*)}{\partial v} \tag{17}$$

and exactly m out of five coefficients λ_i^* of the linear combination have nonzero value. The relationship between the gradient of the minimized function and the gradients of constraint functions leads to the Lagrangian function [3]:

$$L(v, \lambda) = Q(v) - \sum_{i=1}^{5} \lambda_i g_i(v) \tag{18}$$

where vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_5)$ contains the Lagrange multipliers. In Lagrange's method, the search for maximum of $Q(v)$ subject to constraints $g_i(0) = 0$ is translated into the problem of searching for the stationary points of Lagrangian L. The following theorem summarizes the Lagrange method [3]:

Theorem 1 *The necessary condition for function $Q(v)$ to attain an extremum v^* with constraints $g_i(v)$ is that a vector λ^* exists such that the gradient of the Lagrangian function L equals zero: $\nabla L(v^*, \lambda^*) = 0$.*

Note that in order to find the coordinates of the extremum point v^* along with the Lagrange multipliers λ^* using this theorem more than five equations will need to be solved. However, the nature of these equations is not complicated because of the fact that function Q is reducible to a rational function. The first set of five equations originates from $\nabla_\lambda L(v^*, \lambda^*) = \vec{0}$. These have an obvious solution $g(v^*) = \vec{0}$. Such a result could be inferred directly from (8) if all the constraints were known to be satisfied:

$$\begin{aligned} e^{v_1^*} &= 2X_1 + X_3 \\ e^{v_2^*} &= X_2 \\ e^{v_3^*} &= X_2X_4 + X_5(X_1 + X_2 + X_3) \\ e^{v_4^*} &= X_4 \\ e^{v_5^*} &= X_5 \end{aligned} \tag{19}$$

The answer to the question as to which of the constraints should be taken into consideration is in the Lagrange multipliers. They can be found from the second set of five equations originating from $\nabla_v L(v^*, \lambda^*) = \vec{0}$. Although this approach would lead to finding the maximum with the constraints (14), it would by overly complicated by the fact that the constraint equations represent a hypercube in R^5 which is not a differentiable manifold. One of the ways to simplify the task is to realize that the lower and the upper bounds X_{iL} and X_{iU} are just estimates coming from the physical reality. Since they are estimates anyway, then the actual limitations X_{iL} and X_{iU} can be relaxed within a certain tolerance, rather than given exact values. By doing so, the edges and the corners of the hypercube would get blurred within the tolerance limits.

Let $\mu_i = (X_{iU} + X_{iL})/2$ be the coordinates of the center of the hypercube and $\rho_i = (X_{iU} - X_{iL})/2$ be half the lengths of the hypercube sides in each dimension i. The hypersphere inscribed in this hypercube is the set of all zeros of function $S(x) = 1 - \sum_i (x_i - \mu_i)^2/\rho_i^2$. Consider a diffeomorphic deformation of the hypersphere, sufficiently close to the hypercube to fall within the tolerance limits mentioned. Such a deformation can be obtained in many ways using suitable invertible mappings, such as $\tan$ — $\arctan$:

$$G(x) = 1 - \sum_i \left[\frac{1}{\varepsilon}\tan\left(\frac{x_i - \mu_i}{\rho_i}\arctan\varepsilon\right)\right]^2 \tag{20}$$

Note that function $G(x)$ represents the hypersphere as $\varepsilon > 0$ approaches zero: $\lim_{\varepsilon\to 0} G(x) = S(x)$. At the other extreme, as $\varepsilon \to \infty$ function $G(x)$ becomes a representation of the hypercube. Most importantly, as long as ε is finite, G is differentiable and represents the hypercube constraint fairly well for sufficiently large ε. Finding the maximum of $Q(v)$ subject to $G[x(v)]$ requires redefining the Lagrangian:

$$L_G(v, \lambda_G) = Q(v) - \lambda_G G[x(v)] \tag{21}$$

In order to find a point v^* and the Lagrange multiplier λ_G^* the five equations $\nabla_v L(v^*, \lambda_G^*) = \vec{0}$ with one additional equation $\frac{\partial}{\partial\lambda_G} L(v^*, \lambda_G^*) = 0$ need to be solved. To accomplish this task a computer routine for finding roots of functions can be used. Typically such routines use various flavors of the Newton-Raphson method which requires an initial condition $(\hat{v}, \hat{\lambda}_G)$. Moreover, the initial condition should be close enough to the expected solution to ensure convergence. A random initial vector $\hat{v}$ located on the hypersurface such that $G[x(\hat{v})] = 0$ is a good guess. Thus, let φ, θ, ψ, and γ be random numbers from a uniform distribution over the set $(-\pi, \pi)$. The initial vector $\hat{x} = x(\hat{v})$ that solves equation $G(\hat{x})$ can be found from

$$\begin{aligned} \hat{x}_1 &= \mu_1 + \rho_1 \arctan(\varepsilon\cos\varphi)/\arctan\varepsilon \\ \hat{x}_2 &= \mu_2 + \rho_2 \arctan(\varepsilon\sin\varphi\cos\theta)/\arctan\varepsilon \\ \hat{x}_3 &= \mu_3 + \rho_3 \arctan(\varepsilon\sin\varphi\sin\theta\cos\psi)/\arctan\varepsilon \\ \hat{x}_4 &= \mu_4 + \rho_4 \arctan(\varepsilon\sin\varphi\sin\theta\sin\psi\cos\gamma)/\arctan\varepsilon \\ \hat{x}_5 &= \mu_5 + \rho_5 \arctan(\varepsilon\sin\varphi\sin\theta\sin\psi\sin\gamma)/\arctan\varepsilon \end{aligned} \tag{22}$$

The initial value of $\hat{\lambda}_G$ could be any number for which the numeric routine converges. If the routine finds a minimum instead of the maximum of Q, the sign of $\hat{\lambda}_G$ should be changed. The convergence can be improved by lowering the value of ε, however, ε should be large enough to make G close to the hypercube.

Assuming feasible boundaries $(1\,\mathrm{pF}, 1\,\mu\mathrm{F})$ for capacitances and $(1\,\Omega, 1\,\mathrm{M}\Omega)$ for resistances in (6), and setting the constant $\varepsilon = 10000$, the maximum is numerically found to be $Q_{\max} = 0.70709$ at point x^* corresponding to parameters $C_1 = 1\,\mu\mathrm{F}$, $C_2 = 31.75\,\mathrm{pF}$, $C_3 = 29.77\,\mathrm{pF}$, $G_0 = (1\,\Omega)^{-1}$, and $G_2 = (63.02\,\mathrm{k}\Omega)^{-1}$.

IV. Conclusions

Particular details of both procedures for determination of the maximum value of Q will vary from case to case. Nevertheless, this effort will be worthwhile, especially if attempted prior to the construction of the actual circuit. Moreover, the optimum performance of a circuit in terms of selectivity can be determined in this manner.

References

[1] A. Budak, *Passive and Active Network Analysis and Synthesis*. Waveland Press, 1991.

[2] C. H. Edwards, *Advanced Calculus of Several Variables*. London: Academic Press, Inc., 1973.

[3] M. R. Hestenes, *Optimization Theory: the Finite Dimensional Case*. New York: John Wiley & Sons, 1975.