Synthesis of Mechanical Networks: The Inerter

Malcolm C. Smith, *Fellow, IEEE*

***Abstract*—This paper is concerned with the problem of synthesis of (passive) mechanical one-port networks. One of the main contributions of this paper is the introduction of a device, which will be called the inerter, which is the true network dual of the spring. This contrasts with the mass element which, by definition, always has one terminal connected to ground. The inerter allows electrical circuits to be translated over to mechanical ones in a completely analogous way. The inerter need not have large mass. This allows any arbitrary positive-real impedance to be synthesized mechanically using physical components which may be assumed to have small mass compared to other structures to which they may be attached. The possible application of the inerter is considered to a vibration absorption problem, a suspension strut design, and as a simulated mass.**

***Index Terms*—Brune synthesis, Darlington synthesis, electrical–mechanical analogies, mechanical networks, network synthesis, passivity, suspension systems, vibration absorption.**

I. Introduction

THERE is a standard analogy between mechanical and electrical networks in which force (respectively, velocity) corresponds to current (respectively, voltage) and a fixed point in an inertial frame of reference corresponds to electrical ground [9], [26]. In this analogy, the spring (respectively, damper, mass) corresponds to the inductor (respectively, resistor, capacitor). It is well known that the correspondence is perfect in the case of the spring and damper, but there is a restriction in the case of the mass. This restriction is due to the fact that the force–velocity relationship satisfied by the mass, namely Newton's Second Law, relates the acceleration of the mass relative to a fixed point in the inertial frame. Effectively this means that one "terminal" of the mass is the ground and the other "terminal" is the position of the center of mass itself [26, p. 111], [15, pp. 10–15]. Clearly, in the electrical context, it is not required that one terminal of the capacitor is grounded. This means that an electrical circuit may not have a direct spring–mass-damper mechanical analog.

There is a further drawback with the mass element as the analog of the capacitor in the context of *synthesis* of mechanical impedances. Namely, it may be important to assume that the mechanical device associated with the "black-box impedance" to be designed has negligible mass compared to other masses in the system (cf., a suspension strut for a vehicle compared to the sprung and unsprung masses). Clearly this presents a problem if (possibly) large masses may be required for its realization.

It appears that the aforementioned two difficulties have prevented electrical circuit synthesis from being fully exploited for the synthesis of mechanical networks. It seems interesting to ask if these drawbacks are essential ones? It is the purpose of this paper to show that they are not. This will be achieved by introducing a mechanical circuit element, which will be called the *inerter*, which is a genuine two-terminal device equivalent to the electrical capacitor. The device is capable of simple realization, and may be considered to have negligible mass and sufficient linear travel, for modeling purposes, as is commonly assumed for springs and dampers. The inerter allows classical results from electrical circuit synthesis to be carried over exactly to mechanical systems.

Three applications of the inerter idea will be presented. The first is a vibration absorption problem whose classical solution is a tuned spring–mass attached to the main body. It will be shown that the inerter offers an alternative approach which does not require additional elements to be mounted on the main body. The second application is a suspension strut design. Traditional struts employ springs and dampers only, which greatly restricts the available mechanical admittances. In particular, their phase characteristic is always lagging. By considering a general class of third order admittances it will be shown that the use of inerters offers a possibility to reduce oscillation in stiffly sprung suspension systems. The procedures of Brune and Darlington will be employed to obtain network realizations of these admittances. The third application is the use of the inerter to simulate a mass element.

The approach used for the mechanical design problems in this paper owes a debt to the methods of modern control. Firstly, the problems are viewed as an interconnection between a given part of the system (analogous to the plant) and a part to be designed (analogous to the controller). Secondly, the part to be designed is a dynamical element whose admissibility is defined as broadly as possible—passive in the present case (stabilizing for feedback control). The advantage of this viewpoint is that synthesis methods come into play, and that new solutions emerge which would otherwise be missed.

II. Mechanical Networks

A. Classical Network Analogies

Historically, the first analogy to be used between electrical and mechanical systems was the force–voltage analogy, as is readily seen in the early use of the term electromotive force. The alternative force–current analogy is usually attributed to Firestone [9], though it appears to have been independently discovered in [12], [7]. Firestone also introduced the ideas of through and across variables which provide a unifying framework to extend analogies to other contexts, e.g., acoustic, thermal, fluid systems. The reader is referred to [26] for a seminal exposition of this approach (see also [19] and [20]). Interesting historical notes can be found in [22], [18, Ch. 9], [16, Preface].

Manuscript received November 1, 2001; revised April 9, 2002. Recommended by Associate Editor K. Gu. This work was supported in part by the EPSRC.

The author is with the Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, U.K. (e-mail: mcs@eng.cam.ac.uk).

Digital Object Identifier 10.1109/TAC.2002.803532.

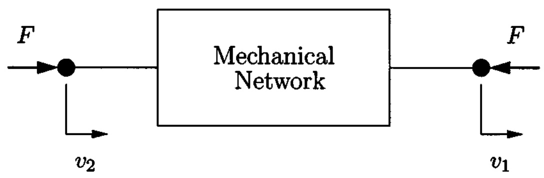

Fig. 1. A free-body diagram of a one-port (two-terminal) mechanical element or network with force–velocity pair (F, v) where $v = v_2 - v_1$.

The subject of dynamical analogies relies strongly on the use of energy ideas, with the product of through and across variables being an instantaneous power. Although there is a sense in which both analogies are valid, the force–current analogy is the one which respects the manner of connection (i.e., series, parallel etc.) so that mechanically and electrically equivalent circuit diagrams are identical as graphs [9], [12], [7]. At a more fundamental level, this arises because the through and across variable concepts allow a direct correspondence between nodes, branches, terminals, and ports in a network [30]. In the closely related bond graph approach to system modeling [23], [16], [17], the use of effort and flow variables, whose product has units of power, normally employs the force–voltage analogy, but this is not intrinsic to that approach [31].

The force–current analogy, described in more detail in Section II-B, is the one preferred here. However, the contribution of the present work is not dependent on which analogy is used. The property of the mass element, that one of its terminals is the ground, is a "restrictive" feature independent of whether its electrical analog is considered to be the capacitor or the inductor. In this sense, the defining property of the inerter is that it is the true mechanical dual of the spring.

B. The Force–Current Analogy

The formal definitions of nodes, branches, elements, etc. in electrical network theory are quite standard and do not need to be repeated here (see [2] for a summary). The analogous but slightly less familiar definitions for mechanical networks will be useful to record below (see [26] for a comprehensive treatment).

A (idealized) *mechanical network* of pure translational type consists of mechanical elements (such as springs, masses, dampers and levers) which are interconnected in a rigid manner [26], [15]. It is usual to restrict the motion to be parallel to a fixed axis and relative to a fixed reference point in an inertial frame called the *ground.* The pair of end-points of the spring and damper are called *nodes* (or *terminals*). For the mass, one terminal is the position of its center of gravity, whilst the other terminal is the ground.

A *port* is a pair of nodes (or terminals) in a mechanical system to which an equal and opposite force F is applied and which experience a relative velocity v. Alternatively, a velocity can be applied which results in a force. Fig. 1 is a free-body diagram of a one-port (two-terminal) mechanical network which illustrates the sign convention whereby a positive F gives a compressive force and a positive $v = v_2 - v_1$ corresponds to the nodes moving together. The product of F and v has units of power and we call (F, v) the force–velocity pair. In general, it is not necessary for either node in a port to be grounded.

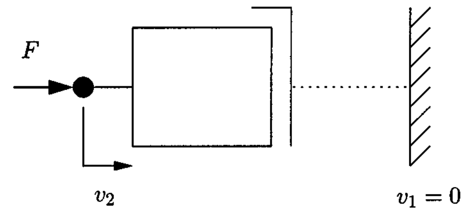

Fig. 2. The standard network symbol for the mass element.

The force–current (sometimes termed *mobility*) analogy between electrical and mechanical networks can be set up by means of the following correspondences:

force	$\leftrightarrow$	current
velocity	$\leftrightarrow$	voltage
mechanical ground	$\leftrightarrow$	electrical ground
spring	$\leftrightarrow$	inductor
damper	$\leftrightarrow$	resistor
kinetic energy	$\leftrightarrow$	electrical energy
potential energy	$\leftrightarrow$	magnetic energy.

The correspondence between mass and capacitor was omitted from the previous list due to the fact that one terminal of the mass element is mechanical ground, which means that the defining equation is analogous to that of the capacitor, but is not as general. This is embodied in the standard network symbol for the mass shown in Fig. 2 where the bracket and dashed line emphasize that v_2 must be measured relative to a nonaccelerating (usually zero velocity) reference.

The force–current analogy goes deeper than the correspondences listed in the previous paragraph because of the concept of through and across variables [9]. In essence, a *through variable* (such as force or current) involves a single measurement point and requires the system to be severed at that point to make the measurement. In contrast, an *across variable* (such as velocity or voltage) can be measured without breaking into the system and the relevant quantity for network analysis is the difference of the variable between two points, even if one point is the ground. A general approach to network analysis based on such a formalism is given in [30]. One consequence is that the methods of mesh- and nodal-analysis can be applied to mechanical networks.

In this paper, we define *impedance* to be the ratio of the across variable to the through variable, which agrees with the usual electrical terminology. For mechanical networks, impedance is then the ratio of velocity to force, which agrees with some books [26, p. 328], but not others which use the force–voltage analogy [15]. We define *admittance* to be the reciprocal of impedance.

C. The Inerter

We define the (ideal) inerter *to be a mechanical two-node (two-terminal), one-port device with the property that the equal and opposite force applied at the nodes is proportional to the relative acceleration between the nodes.* That is, $F = b(\dot{v}_2 - \dot{v}_1)$ in the notation of Fig. 1. The constant of proportionality b is called the *inertance* and has units of kilograms. The stored energy in the inerter is equal to $(1/2)b(v_2 - v_1)^2$.

Naturally, such a definition is vacuous unless mechanical devices can be constructed which approximate the behavior of the ideal inerter. To be useful, such devices also need to satisfy certain practical conditions which we list as follows.

R1) The device should be capable of having a small mass, independent of the required value of inertance.
R2) There should be no need to attach any point of the physical device to the mechanical ground.
R3) The device should have a finite linear travel which is specifiable, and the device should be subject to reasonable constraints on its overall dimension.
R4) The device should function adequately in any spatial orientation and motion.

Condition R2) is necessary if the inerter is to be incorporated in a free-standing device which may not easily be connected to a fixed point in an inertial frame, e.g., a suspension strut which is connected between a vehicle body and wheel hub. We mention that conditions of the above type hold for the ordinary spring and damper.

The aforementioned realizability conditions can indeed be satisfied by a mechanical device which is easy to construct. A simple approach is to take a plunger sliding in a cylinder which drives a flywheel through a rack, pinion, and gears (see Fig. 3). Note that such a device does not have the limitation that one of the terminals be grounded, i.e., attached to a fixed point in an inertial frame. To approximately model the dynamics of the device of Fig. 3, let r_1 be the radius of the rack pinion, r_2 the radius of the gear wheel, r_3 the radius of the flywheel pinion, γ the radius of gyration of the flywheel, m the mass of the flywheel, and assume the mass of all other components is negligible. Assuming $v_1 = 0$ we can check that the following relation holds:

$$F = (m\alpha_1^2\alpha_2^2)\dot{v} \tag{1}$$

where $\alpha_1 = \gamma/r_3$ and $\alpha_2 = r_2/r_1$. If $v_1 \neq 0$ the direct inertial effect of the flywheel mass comes into play, but this will only change (1) by a small proportion providing $\alpha_1^2\alpha_2^2$ is large. To a first approximation, such an effect can be neglected, as is commonly done for springs and dampers. Note that even with relatively modest ratios $\alpha_1 = \alpha_2 = 3$ the inertance is a factor of 81 times the mass. It is clearly feasible to introduce additional gearing; an extra gear wheel and pinion with ratio α_3 will multiply the inertance by a factor α_3^2. Increasing the gearing ratios also increases internal forces in the device and the flywheel angular velocity (the latter is given by $r_2 r_1^{-1} r_3^{-1}(\dot{v}_2 - \dot{v}_1)$ in the above model) which places higher demands in manufacture, but these are practical concerns and not fundamental limits. In principle, it is feasible to keep the mass of the device small in an absolute sense, and compared to the inertance of the device. Indeed a simple prototype inerter has been made which has an inertance to mass ratio of about 300.[1] The remaining conditions R2)–R4) are also satisfied by the realization of Fig. 3. In the case of gyroscopic effects being an issue under R4), a system of counter-rotating flywheels could be introduced. It seems reasonable to conclude that such a device can be regarded as approximating the ideal inerter in the same sense that real springs, dampers, inductors, resistors, and capacitors approximate their mathematical ideals.

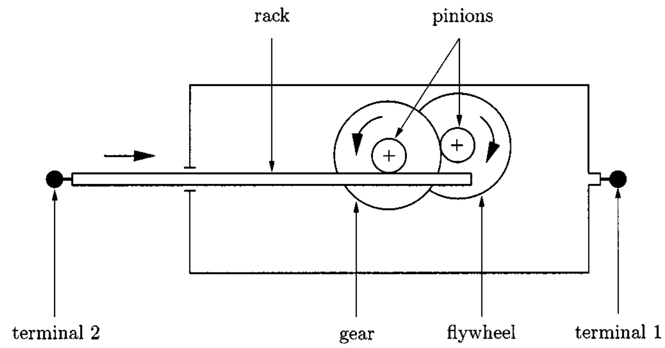

Fig. 3. Schematic of a mechanical model of an inerter.

It is useful to discuss two references on mechanical networks, which give some hint toward the inerter idea, in order to highlight the new contribution here. We first mention [26, p. 234] which describes a procedure whereby an electrical circuit is first modified by the insertion of ideal one-to-one transformers so that all capacitors then have one terminal grounded. This then allows a mechanical circuit to be constructed with levers, which has similar dynamic properties to the electrical one while not being properly analogous from a circuit point of view. Condition R1) is not discussed in [26] though it seems that this could be addressed by adjusting the transformer ratios to reduce the absolute values of the masses required (with transformers then being needed for all capacitors), however, R3) might then be a problem. Another difficulty with this approach is with R2) since a pair of terminals of the transformer need to be connected to the mass and the ground.

Second, we highlight the paper of Schönfeld [24], which is principally concerned with the treatment of hydraulic systems as distinct from mechanical systems and the interpretation of acoustic systems as mixed mechanical–hydraulic systems, a work which appears to have been unfairly neglected. In connection with mechanical–electrical analogies, the possibility of a biterminal mechanical inertance is mentioned. The idea is essentially to place a mass at the end of a lever, connected with links to the two terminals, while increasing the lever length and decreasing the value of mass arbitrarily but in fixed ratio [24, Fig. 12(d)]. Although this in principle deals with R1) and R2), there is a problem with R3) due to the large lever length required or the vanishing small available travel. A variant on this idea [24, Fig. 12(e)] has similar difficulties as well as a problem with R4). It is perhaps the obvious limitations of these devices that have prevented the observation from being developed or formalized.

In the light of the previous definition of the ideal inerter, it may sometimes be an advantage to reinterpret combinations of system elements as acting like an inerter. For example, in [17, Problem 4.18] two masses are connected together by means of a lever arrangement (interpreted as a 2-port transformer connected to a 1-port inertia element in the bond graph formalism). If this system is linearized for small displacements then the behavior is the same as if an inerter were connected between the two masses. Of course, such an arrangement has problems with

[1]Patent pending.

R3). Indeed, if large values of inertance were required for a moderate amount of travel then the lever lengths and ratios would be impractical.

A table of the circuit symbols of the six basic electrical and mechanical elements, with the newly introduced inerter replacing the mass, is shown in Fig. 4. The symbol chosen for the inerter represents a flywheel.

D. Classical Network Synthesis

The introduction of the inerter mechanical element, and the use of the force–current analogy, allows a classical theorem on synthesis of electrical one-ports in terms of resistors, capacitors and inductors to be translated over directly into the mechanical context. We will now restate the relevant definitions and results in mechanical terms.

Consider a mechanical one-port network as shown in Fig. 1 with force–velocity pair (F, v). The network is defined to be *passive* [21, p. 26], [1, p. 21] if for all admissible v, F which are square integrable on $(-\infty, T]$

$$\int_{-\infty}^{T} F(t)v(t)\,dt \geq 0. \tag{2}$$

The quantity on the left-hand side of (2) has the interpretation of the total energy delivered to the network up to time T. Thus, a passive network cannot deliver energy to the environment.

Theorem 1 [21, Chs. 4, 5], [1, Th. 2.7.1, 2]: Consider a one-port mechanical network for which the impedance $Z(s)$ exists and is real-rational. The network is passive if and only if one of the following two equivalent conditions is satisfied.

1) $Z(s)$ is analytic and $Z(s) + Z(s)^* \geq 0$ in $\mathrm{Re}(s) > 0$.
2) $Z(s)$ is analytic in $\mathrm{Re}(s) > 0$, $Z(j\omega) + Z(j\omega)^* \geq 0$ for all ω, at which $Z(j\omega)$ is finite, and any poles of $Z(s)$ on the imaginary axis or at infinity are simple and have a positive residue.

In the aforementioned theorem, * denotes complex conjugation. A pole is said to be *simple* if it has multiplicity one. The residue of a simple pole of $Z(s)$ at p_0 is equal to $\lim_{s\to p_0}(s - p_0)Z(s)$. Poles and zeros of $Z(s)$ at $s = \infty$ can be defined as the poles and zeros of $Z(s^{-1})$ at $s = 0$. Thus the residue of a simple pole at $s = \infty$ is equal to $\lim_{s\to\infty} Z(s)/s$.

Any real-rational function $Z(s)$ satisfying 1) or 2) in Theorem 1 is called *positive real*. Theorem 1 also holds with $Z(s)$ replaced by the admittance $Y(s)$.

Theorem 2: Consider any real-rational function $Z(s)$ which is positive real. There exists a one-port mechanical network whose impedance equals $Z(s)$ which consists of a finite interconnection of springs, dampers, and inerters.

Theorem 2 is also valid with $Z(s)$ replaced by the admittance $Y(s)$. This theorem represents one of the key results of classical electrical network synthesis, translated directly into mechanical terms. The first proof of a result of this type was given in [4], which shows that any real-rational positive-real function could be realized as the driving-point impedance of an electrical network consisting of resistors, capacitors, inductors, and transformers. The method involves a sequence of steps to successively reduce the degree of the positive-real function by extraction of imaginary axis poles and zeros and

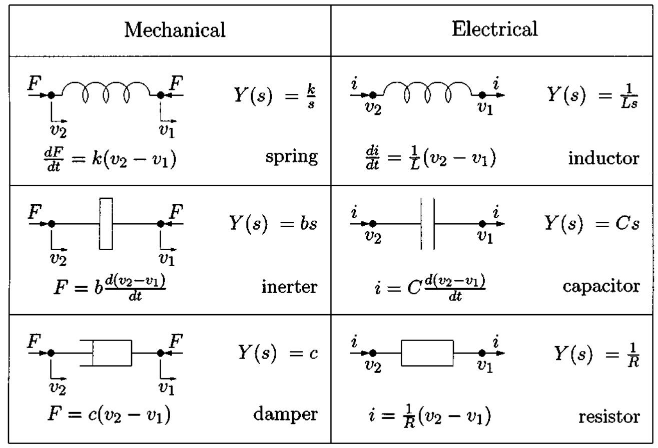

Fig. 4. Circuit symbols and correspondences with defining equations and admittance $Y(s)$.

subtraction of resistive and reactive elements [11, Ch. 9.4], [4]. A classical alternative procedure due to Darlington [5] realizes the impedance as a lossless two-port network terminated in a single resistance. The possibility of achieving the synthesis without the use of transformers was first established by Bott and Duffin [3]. See [11, Ch. 10] and [2, pp. 269–274] for a description of this and related methods, and [6] for a historical perspective. It is these procedures which provide the proof for Theorem 2.

III. Vibration Absorption

A. Problem Statement

Suppose we wish to connect a mass M to a structure so that steady sinusoidal vibrations of the structure at a constant frequency ω_0 do not disturb the mass. The problem is posed as in Fig. 5 where the mass is connected to the structure by a device whose mechanical admittance is $Q(s)$. The mass may be subjected to a force F_L and the displacement of the mass and the structure are x and z, respectively. We seek to design and realize a positive-real $Q(s)$ so that if $z = \sin(\omega_0 t)$ then $x(t) \to 0$ as $t \to \infty$.

The equation of motion for the mass M in the Laplace transformed domain is:

$$Ms^2\hat{x} = \hat{F}_L + sQ(s)(\hat{z} - \hat{x})$$

whence

$$\hat{x} = \frac{1}{Ms^2 + sQ(s)}\hat{F}_L + \frac{Q(s)}{Ms + Q(s)}\hat{z}$$

where $\hat{}$ denotes Laplace transform. It is evident that the mass will be impervious to a steady sinusoidal disturbance at z providing $Q(s)/(Ms + Q(s))$ has a zero at $s = j\omega_0$, and that this will hold providing $Q(s)$ has a zero at $s = j\omega_0$.

B. Approach Using Inerter

Let us seek an admittance of the form

$$Q(s) = c\,\frac{(s+a)(s^2+\omega_0^2)}{s(s^2+b_1 s+b_2)} \tag{3}$$

with c, a, b_1, $b_2 > 0$. The reasoning for this choice of $Q(s)$ is as follows. If the quadratic factors in the numerator and denominator are removed then the admittance reduces to that of a spring and damper in parallel. The factor $s^2 + \omega_0^2$ gives the required zero at $s = j\omega_0$ in $Q(s)$ and the quadratic factor in the denominator allows $Q(s)$ to approximate the behavior of the spring and damper for large s and for small s.

We require that $Q(s)$ is positive real so it can be realized passively. Consider the positive real factor $Q_1(s) = (s^2 + \omega_0^2)/s$ in $Q(s)$. Note that $Q_1(j\omega)$ is purely imaginary, with a positive sign if $0 < \omega < \omega_0$ and a negative sign if $\omega > \omega_0$. Considering the behavior of $Q(j\omega)$ near $\omega = \omega_0$ it is evident that $\mathrm{Re}(Q(j\omega)) \geq 0$ for all ω only if $(s + a)/(s^2 + b_1 s + b_2)$ is real when $s = j\omega_0$. The latter condition holds providing that

$$b_2 = \omega_0^2 + a b_1. \quad (4)$$

It turns out that (4) is also sufficient for $Q(s)$ in (3) to be positive real when c, a, b_1, b_2, > 0. Rather than verify this directly, we will now consider how $Q(s)$ can be realized.

A standard first step in synthesizing a positive real function is to remove any imaginary axis poles and zeros [4], [11, Ch. 9.4]. For the function in (3) it turns out to be simplest to remove first the zeros at $s = \pm j\omega_0$ by considering $Q^{-1}(s)$. We obtain

$$\begin{aligned} Q^{-1}(s) &= \frac{s(s^2 + b_1 s + b_2)}{c(s+a)(s^2+\omega_0^2)} \\ &= \frac{b_1}{c} \cdot \frac{s}{s^2+\omega_0^2} + \frac{1}{c} \cdot \frac{s}{s+a} \\ &=: Z_1(s) + Z_2(s) \end{aligned} \quad (5)$$

using (4). Equation (5) gives a preliminary decomposition of $Q(s)$ as a series connection of two network elements with mechanical impedances $Z_1(s)$ and $Z_2(s)$ respectively. The first of these elements has an admittance given by

$$Z_1^{-1}(s) = \frac{c}{b_1} s + \frac{c\omega_0^2}{b_1} \cdot \frac{1}{s}$$

which represents a parallel combination of an inerter with constant c/b_1 and a spring with constant $c\omega_0^2/b_1$. The second element, called the *minimum reactive* part in electrical networks [11, Ch. 8.1], has an admittance given by

$$Z_2^{-1}(s) = c + \frac{ca}{s}$$

which represents a parallel combination of a damper with constant c and spring of constant ca. Writing $k_1 = c\omega_0^2/b_1$ and $k_2 = ca$ we therefore obtain the realization of $Q(s)$ as shown in Fig. 6.

We remark that the new feature in the admittance $Q(s)$ is the presence of the parallel combination of the inerter and spring. This is, in fact, a tuned linear oscillator with natural frequency of oscillation ω_0.

C. Comparison with Conventional Approach

It is interesting to compare the solution obtained in Fig. 6 with the more conventional approach shown in Fig. 7 where the vibration absorber consists of a tuned spring-mass system

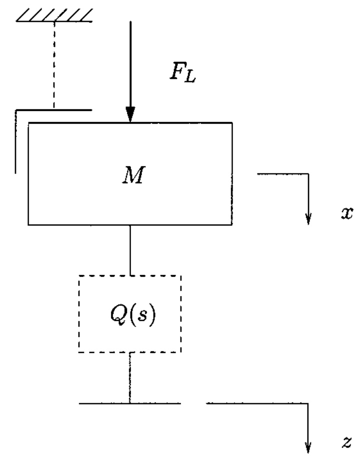

Fig. 5. Vibration absorption problem.

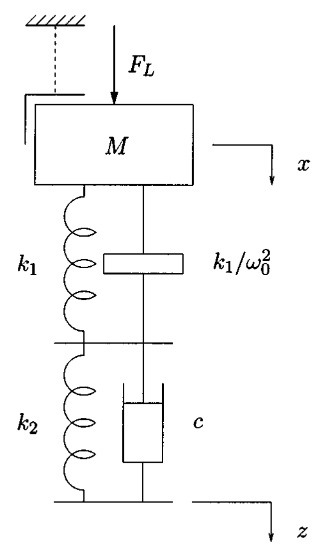

Fig. 6. Realization of $Q(s)$.

attached to the mass M (see [8]). In the Laplace transformed domain, the equations of motion are

$$\begin{aligned} ms^2\hat{y} &= k_4(\hat{x} - \hat{y}) \\ Ms^2\hat{x} &= \hat{F}_L + k_4(\hat{y} - \hat{x}) + (k_3 + c_1 s)(\hat{z} - \hat{x}). \end{aligned}$$

Solving for $\hat{x}$ and $\hat{y}$ gives

$$\begin{aligned} \hat{x} &= (1 + s^2/\omega_0^2)\hat{y} \\ \hat{y} &= \frac{(k_3 + c_1 s)\hat{z} + \hat{F}_L}{(Ms^2 + c_1 s + (k_3 + k_4))(1 + s^2/\omega_0^2) - k_4}. \end{aligned}$$

Thus, when $F_L = 0$, the mass M has zero steady-state amplitude in response to a sinusoidal disturbance at z of unit amplitude and frequency ω_0 (which is the desired vibration absorption property) while the steady-state amplitude of the attached mass m is

$$\frac{\sqrt{k_3^2 + c_1^2\omega_0^2}}{k_4}.$$

It is evident that the amplitude of oscillation of the mass m may be large if k_4 is small compared to $\sqrt{k_3^2 + c_1^2\omega_0^2}$. Thus,

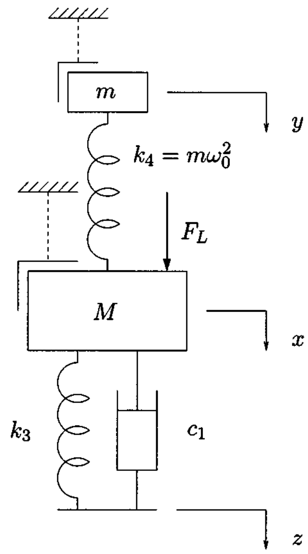

Fig. 7. Conventional vibration absorber.

in practice, m and k_4 will need to be sufficiently large to avoid excessive oscillation in m. This may be a disadvantage if it is undesirable to mount too much additional mass on M.

By contrast, in the solution using the inerter in Fig. 6, the "compensation" for the oscillation of the structure to which M is attached occurs entirely within the device implementing the admittance $Q(s)$. The desired effect is achieved for any value of $k_1 > 0$. What then is the role of k_1 in the performance of this system? Clearly, the choice of k_1 plays a role in the transient response and the response to other disturbances of the structure. Also, the response to loads F_L depends on k_1. In particular, the static spring stiffness under loads F_L equals $k_1k_2/(k_1+k_2)$, which suggests that k_1 should not be too small compared to k_2 for the mass M to be well supported. Unlike the vibration absorber of Fig. 7 there is no objection to increasing k_1 on the grounds of adding extra mass to the structure.

It is instructive to compare the dynamic response of the two solutions to the vibration absorption problem. Let $T^{(1)}_{\hat{z}\to\hat{x}}$ (respectively, $T^{(2)}_{\hat{z}\to\hat{x}}$) denote the transfer function from $\hat{z}$ to $\hat{x}$ for the solution of Fig. 6 (respectively, Fig. 7). Then

$$T^{(1)}_{\hat{z}\to\hat{x}} = \frac{(cs+k_2)(s^2+\omega_0^2)}{d_1(s)}$$

and

$$T^{(2)}_{\hat{z}\to\hat{x}} = \frac{(c_1s+k_3)(s^2+\omega_0^2)}{d_2(s)}$$

where

$$\begin{aligned} d_1(s) =& Ms^4 + c(1+M\omega_0^2k_1^{-1})s^3 \\ &+ (M\omega_0^2(1+k_2k_1^{-1})+k_2)s^2 + c\omega_0^2 s + k_2\omega_0^2 \\ d_2(s) =& Ms^4 + c_1s^3 + (M\omega_0^2+k_3+k_4)s^2 + c_1\omega_0^2 s + k_3\omega_0^2. \end{aligned}$$

Clearly, $T^{(1)}_{\hat{z}\to\hat{x}}(j\omega_0) = T^{(2)}_{\hat{z}\to\hat{x}}(j\omega_0) = 0$, as was required of each approach. The two transfer functions have a similar form, and behave similarly in the limit as $k_1 \to \infty$ (respectively $k_4 \to 0$) as we will see below. For any $\omega \neq \omega_0$ it is straightforward to show that

$$T^{(1)}_{\hat{z}\to\hat{x}}(j\omega) \to \frac{cj\omega + k_2}{-M\omega^2 + cj\omega + k_2} \tag{6}$$

as $k_1 \to \infty$. The transfer function on the right-hand side of (6) is the one obtained when the spring–inerter combination in Fig. 6 is removed. Because of the pointwise convergence described by (6), $T^{(1)}_{\hat{z}\to\hat{x}}$ has the appearance of a "notch" filter (with zero at $\omega=\omega_0$) with an increasingly narrow notch as k_1 becomes large. It is evident that the vibration absorption property is more sensitive to variations in the disturbing frequency ω_0 as k_1 becomes large. That is, if the structure is designed for a disturbance frequency ω_0, but in reality the frequency is $\omega_0+\epsilon$ where ϵ is small, then the resulting disturbance attenuation will be ineffective if k_1 has too large a value. Similar considerations apply for the conventional approach of Fig. 7. For any $\omega \neq \omega_0$

$$T^{(2)}_{\hat{z}\to\hat{x}}(jw) \to \frac{c_1jw + k_3}{-M\omega^2 + c_1j\omega + k_3}$$

as $k_4 \to 0$. Again, as k_4 becomes smaller, there is an increasingly narrow notch at $\omega = \omega_0$, and the vibration absorption property becomes more sensitive to variations in the disturbing frequency.

It is also useful to consider the response of x to load disturbances F_L. These take the form

$$T^{(1)}_{\hat{F}_L\to\hat{x}} = \frac{s^2 + c\omega_0^2k_1^{-1}s + \omega_0^2(1+k_2k_1^{-1})}{d_1(s)}$$

and

$$T^{(2)}_{\hat{F}_L\to\hat{x}} = \frac{s^2+\omega_0^2}{d_2(s)}$$

for the two approaches. We can check that for any $\omega \neq \omega_0$

$$T^{(1)}_{\hat{F}_L\to\hat{x}}(j\omega) \to \frac{1}{-M\omega^2 + cj\omega + k_2}$$

as $k_1 \to \infty$. A similar result holds for $T^{(2)}_{\hat{F}_L\to\hat{x}}$ as $k_4 \to 0$ but with c, k_2 replaced by c_1, k_3. We also obtain $T^{(1)}_{\hat{F}_L\to\hat{x}}(j\omega_0) = -1/M\omega_0^2$ and $T^{(2)}_{\hat{F}_L\to\hat{x}}(j\omega_0) = 0$. Thus, the two solutions differ in their response to sinusoidal load disturbances of frequency ω_0. Although the load disturbances response is not a primary consideration in this problem, the effect of these constraints may need to be considered in the choice of parameters. For example, if $1/M\omega_0^2$ is significantly larger than $1/|-M\omega_0^2 + cj\omega_0 + k_2|$, then the dynamic load response in the inerter approach may not be satisfactory. Similar considerations apply for the conventional approach of Fig. 7.

To conclude this discussion, we can say that the inerter offers an interesting alternative solution to a standard vibration absorption problem. The dynamic response properties of the two solutions are broadly similar, as are the asymptotic properties as the additional mass or inertance becomes small or large. The inerter approach has a potential advantage in that there is no need to mount additional mass on M and to be concerned about possible limits of travel of this additional mass.

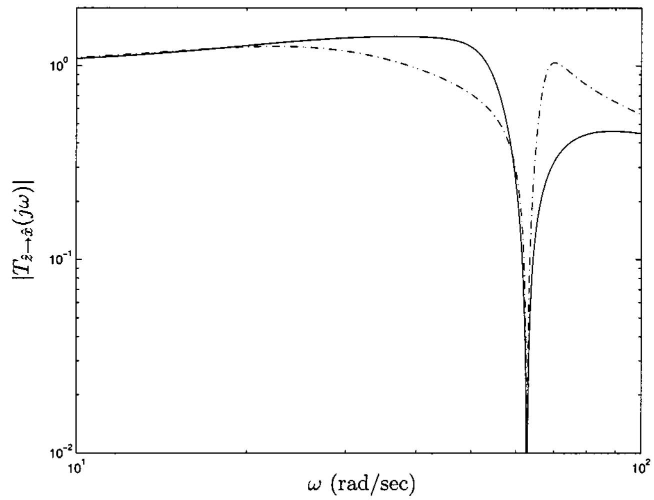

Fig. 8. Frequency responses $T_{\hat{z}\to\hat{x}}$: inerter (–), conventional (–·).

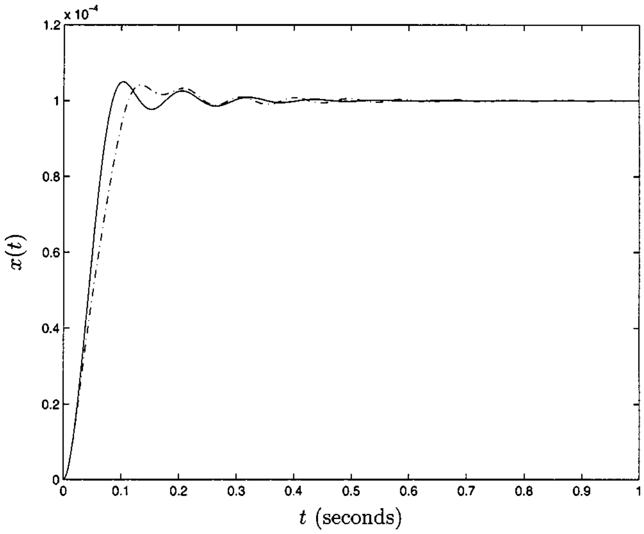

Fig. 9. Response of x to a unit step at F_L: inerter (–), conventional (–·).

D. Numerical Example

Consider the problem as posed in Fig. 5 with $M = 10$ kg and $\omega_0 = 10$ Hz. Suppose it is required that a constant load $F_L = 10$ N produces a deflection at x of at most 1 mm, so that the static spring stiffness is at least 10^4 Nm^{-1}. If $Q(s)$ is chosen to be a spring k in parallel with a damper c, then in Fig. 5

$$T_{\hat{z}\to\hat{x}} = \frac{cs+k}{Ms^2+cs+k} = \frac{2\zeta s/\omega_n + 1}{s^2/\omega_n^2 + 2\zeta s/\omega_n + 1}$$

where $\omega_n^2 = k/M$, $\zeta = c/2\sqrt{kM}$. Setting $k = 10^4$ Nm^{-1} gives $|T_{\hat{z}\to\hat{x}}(j\omega_0)| = 0.828$ when $\zeta = 1$ and $|T_{\hat{z}\to\hat{x}}(j\omega_0)| \to 0.339$ as $\zeta \to 0$. Thus, even in the limit as the damping ratio ζ vanishes (which is likely to be unacceptably oscillatory in any case), the maximal reduction in amplitude is to 34% of the disturbance amplitude. For critical damping the reduction is only to 83%. Evidently, the ordinary spring–damper arrangement is unlikely to provide an acceptable solution for this problem.

Let us begin with the conventional approach of Fig. 7. As m becomes smaller, the "notch" in the frequency response becomes increasingly narrow. Also, we can observe an oscillatory component in the time response which is hard to dampen by adjusting c_1. There is clearly a practical limit to how large m can be. Let us choose the parameters $m = 2$ kg, $k_4 = m\omega_0^2 = 7.896\times10^3$ Nm^{-1}, $k_3 = k = 10^4$ Nm^{-1} and $c_1 = 1.6\sqrt{kM} = 506.0$ Nsm^{-1}. The resulting frequency response $T^{(2)}_{\hat{z}\to\hat{x}}$ is shown in Fig. 8 and the step response of $T^{(2)}_{\hat{F}_L\to\hat{x}}$ is shown in Fig. 9. The steady-state amplitude of m equals $\sqrt{k_3^2 + c_1^2\omega_0^2}/k_4 = 4.22$ times the amplitude of the sinusoidal disturbance at z.

Turning to the approach of Fig. 6 using the inerter. To achieve a static stiffness of $k = 10^4$ Nm^{-1} we need to choose $k_1^{-1} + k_2^{-1} = k^{-1}$. At either extreme for k_1 (close to k or tending toward infinity), we can again observe an oscillatory component in the time response which is hard to dampen by adjusting c. The following parameter choices give a reasonably wide notch and moderate overall damping: $k_1 = 9\times10^4$ Nm^{-1}, $k_2 = (9/8)\times10^4$ Nm^{-1}, $c = 1.7\sqrt{kM} = 537.6$ Nsm^{-1}. This requires an inerter of inertance 22.80 kg. The resulting frequency response $T^{(1)}_{\hat{z}\to\hat{x}}$ is shown in Fig. 8 and the step response of $T^{(1)}_{\hat{F}_L\to\hat{x}}$ is shown in Fig. 9.

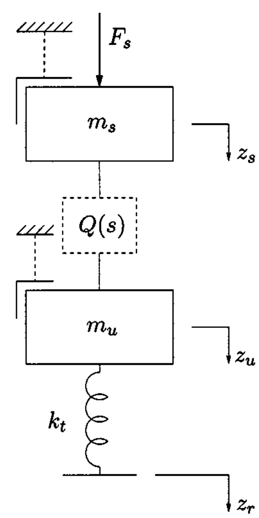

Fig. 10. Quarter-car vehicle model.

IV. Vehicle Suspensions

A. Quarter-Car Model

An elementary model to consider the theory of active and passive suspensions is the quarter-car of Fig. 10 (see, e.g., [25] and [14]) consisting of the sprung mass m_s, the unsprung mass m_u and a tyre with spring stiffness k_t. The *suspension strut* provides an equal and opposite force on the sprung and unsprung masses and is assumed here to be a passive mechanical admittance $Q(s)$ which has negligible mass. The equations of motion in the Laplace transformed domain are

$$m_s s^2 \hat{z}_s = \hat{F}_s - sQ(s)(\hat{z}_s - \hat{z}_u) \tag{7}$$
$$m_u s^2 \hat{z}_u = sQ(s)(\hat{z}_s - \hat{z}_u) + k_t(\hat{z}_r - \hat{z}_u). \tag{8}$$

Using the force–current analogy, the quarter-car model has an electrical analog as shown in Fig. 11, with the two masses becoming grounded capacitors and the two external inputs F_s (loads on the sprung mass) and $\dot{z}_r$ (road undulations modeled as a velocity source) becoming current and voltage sources, respectively.

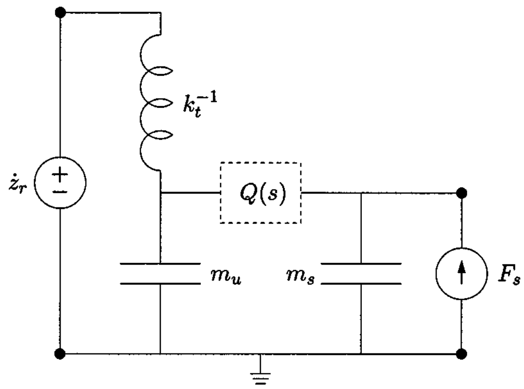

Fig. 11. Equivalent electrical circuit for quarter-car model.

B. Suspension Struts

A fully active suspension allows a much greater design freedom than the traditional suspension struts [27], [29], but there are drawbacks in terms of expense and complexity. Currently, passive suspension struts make use only of springs and dampers. In electrical terms this corresponds to circuits comprising inductors and resistors only. The driving-point impedance or admittance of such circuits is quite limited compared to those using capacitors as well, as is shown by the following result which is translated directly from its electrical equivalent [11, pp. 58–64].

Theorem 3: Consider any one-port mechanical network which consists of a finite interconnection of springs and dampers. If its driving-point admittance exists then it is a (real-rational) positive real function $Y(s)$ with the following properties: all its poles and zeros are simple and alternate on the negative real axis with a pole, possibly at the origin, being the rightmost of these.

For convenience we call a function satisfying the conditions of Theorem 3 a *spring-damper (SD) admittance* (in electrical networks the term RL admittance is used). Any SD admittance $Y(s)$ must satisfy the following two conditions:

$$-90^\circ \leq \arg(Y(j\omega)) \leq 0^\circ \tag{9}$$

$$-20 \text{ dB} \leq \text{Bode-slope}(|Y(j\omega)|) \leq 0 \text{ dB} \tag{10}$$

for $\omega \geq 0$. This follows by considering the contribution of each pole and zero in turn, which gives: $\arg(Y(j\omega)) = -\phi_0 + \phi_1 - \phi_2 + \phi_3 - \phi_4 + \cdots$ where $90^\circ \geq \phi_0 \geq \phi_1 \geq \phi_2 \geq \cdots \geq 0^\circ$. A similar argument proves (10). These conditions on an SD admittance do not apply to a general positive-real admittance, which may exhibit phase lead and has no fundamental restrictions on the local Bode-slope. It seems clear that there could be significant advantages in optimizing the performance of passive suspension systems over the class of positive-real functions. The use of inerters as well as springs and dampers provides the means to do this.

C. Low Degree Positive Real Admittances

It is a general result of network synthesis [11, pp. 127–130] that any SD admittance $Y(s)$ can be realized as in Fig. 12, where n is the number of zeros of $Y(s)$. Even if transformers (levers) are allowed in addition to springs and dampers the class of achievable admittances is still the same as that given by Theorem 3 (see [28]). Thus, the most general SD admittance with (positive-static stiffness) using springs and one damper is given by

$$Y_1(s) = k\,\frac{(T_2 s + 1)}{s(T_1 s + 1)} \tag{11}$$

where $T_2 > T_1 > 0$ and $k > 0$, while the most general form with two dampers is

$$Y_2(s) = k\frac{(T_4 s + 1)(T_6 s + 1)}{s(T_3 s + 1)(T_5 s + 1)} \tag{12}$$

where $T_6 > T_5 > T_4 > T_3 > 0$ and $k > 0$. To investigate the possible benefits that inerters might provide let us consider the class of *arbitrary* positive real mechanical admittances of the same order as Y_2.

Theorem 4: Consider the real-rational function

$$Y(s) = k\frac{a_0 s^2 + a_1 s + 1}{s(d_0 s^2 + d_1 s + 1)} \tag{13}$$

where $d_0, d_1 \geq 0$ and $k > 0$. Then, $Y(s)$ is positive real if only if the following three inequalities hold:

$$\beta_1 := a_0 d_1 - a_1 d_0 \geq 0 \tag{14}$$

$$\beta_2 := a_0 - d_0 \geq 0 \tag{15}$$

$$\beta_3 := a_1 - d_1 \geq 0. \tag{16}$$

Furthermore, $Y(s)$ is an SD admittance of McMillan degree three if and only if $d_0 > 0$, $d_1 > 0$, $\beta_1 > 0$ and the following inequality holds:

$$\beta_4 := \beta_2^2 - \beta_1\beta_3 < 0. \tag{17}$$

(These last four inequalities together imply $\beta_2 > 0$ and $\beta_3 > 0$).

Proof: Assume $Y(s)$ is positive real. We can calculate that

$$\text{Re}(Y(j\omega)) = k\,\frac{a_1 - d_1 + \omega^2(a_0 d_1 - a_1 d_0)}{(d_1\omega)^2 + (1 - d_0\omega^2)^2}. \tag{18}$$

By considering the behavior near $\omega = 0$ and $\omega = \infty$, we conclude that (14) and (16) must hold. Now, if $d_1 > 0$, then

$$a_0 - d_0 = \frac{1}{d_1}\,(a_0 d_1 - a_1 d_0) + \frac{d_0}{d_1}\,(a_1 - d_1)$$

which means that (15) must hold. If $d_0 = 0$ then (15) must again hold since $a_0 < 0$ implies $Y(s)$ has zeros with positive real part which contradicts the positive realness assumption. So let us consider the remaining case where $d_1 = 0$ and $d_0 > 0$. In this case, a pole occurs on the imaginary axis at $s = j/\sqrt{d_0}$. From (14), it follows that $a_1 = 0$, whence the residue at the pole is equal to $k(a_0 - d_0)/(2d_0)$. Since this must be nonnegative this again establishes (15), completing all cases.

For the converse direction, $\text{Re}(Y(j\omega)) \geq 0$ from (18) so we need only check the residue conditions for any imaginary axis poles. The pole at $s = 0$ has residue $k > 0$. A pole at $s = \infty$ can occur if $d_0 = d_1 = 0$ in which case the residue is $a_0 \geq 0$. If $d_1 = 0$ and $d_0 > 0$ then a pole occurs on the imaginary axis at $s = j/\sqrt{d_0}$. Again, $a_1 = 0$ from (14) and the residue at the pole is equal to $k(a_0 - d_0)/(2d_0)$ which is nonnegative. This proves positive realness of $Y(s)$.

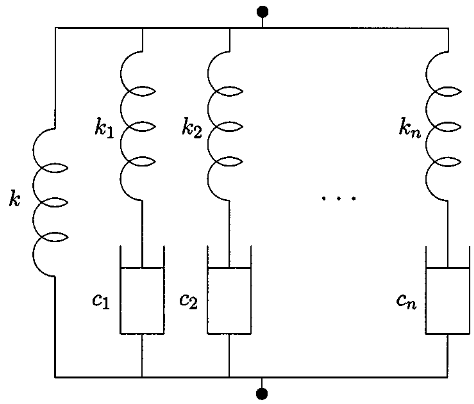

Fig. 12. Realization of a general SD admittance.

Turning to the final claim, $Y(s)$ is an SD admittance of McMillan degree 3 if and only if $Y(s)$ is positive real and satisfies the pole-zero interlacing property of Theorem 3 (with strict inequalities). Using [10, Ch. XV, Th. 11] the interlacing property holds (with strict inequalities) if and only if $d_0s^5 + a_0s^4 + d_1s^3 + a_1s^2 + s + 1$ is Hurwitz. Using the Liénard–Chipart criterion [10, Ch. XV, Th. 13] the latter holds if the only if the following inequalities hold: $d_0 > 0$, $d_1 > 0$

$$\det\begin{bmatrix} a_0 & a_1 \\ d_0 & d_1 \end{bmatrix} > 0, \qquad \det\begin{bmatrix} a_0 & a_1 & 1 & 0 \\ d_0 & d_1 & 1 & 0 \\ 0 & a_0 & a_1 & 1 \\ 0 & d_0 & d_1 & 1 \end{bmatrix} > 0. \quad (19)$$

We can check that the two determinants in (19) are equal to β_1 and $-\beta_4$. It can also be seen that (17) implies $\beta_3 > 0$, whence $\beta_2 > 0$ follows from $\beta_1 > 0$. △

Before studying the possible benefits of the admittance (13), let us consider how it could be realized.

D. Realizations Using Brune Synthesis

The synthesis of *general* positive-real functions cannot be achieved with such a simple canonical form as Fig. 12 and requires the more sophisticated procedures of Section II-D. For the realization of the admittance (13) we can assume without loss of generality that

A1) $\beta_2 > 0$;
A2) $\beta_4 > 0$.

We can justify this as follows. If $\beta_2 = 0$, then either $a_0 = d_0 = 0$ or $\beta_3 = 0$ [by combining (14) and (16)], which in both cases leads to a loss of McMillan degree and the possibility of realization in the form of Fig. 12. If $\beta_4 < 0$, then (17) implies $\beta_1 > 0, \beta_3 > 0$ and (14) then implies $d_1 > 0$. This is sufficient for the pole-zero interlacing property of Theorem 3 to hold (the case of $d_0 = 0$ can be checked as in the proof of Theorem 4) so that realization in the form of Fig. 12 is again possible. Now, if $\beta_4 = 0$ and $\beta_2 > 0$ (which implies that $\beta_1 > 0$) then we can check that $-\beta_2/\beta_1$ is a common pole-zero pair in (13) so again realization is possible in the form of Fig. 12.

The first step in the Brune procedure [4, Ch. 9.4] is to remove the imaginary axis poles and zeros. We write

$$Y(s) = k\frac{a_0s^2 + a_1s + 1}{s(d_0s^2 + d_1s + 1)} = \frac{k}{s} + \frac{k(\beta_2s + \beta_3)}{d_0s^2 + d_1s + 1} \quad (20)$$

and then

$$\frac{d_0s^2 + d_1s + 1}{k(\beta_2s + \beta_3)} = \frac{d_0}{\beta_2k}\,s + \frac{\beta_1s + \beta_2}{\beta_2k(\beta_2s + \beta_3)} \quad (21)$$

$$= \frac{d_0}{\beta_2k}\,s + \frac{\beta_1}{\beta_2^2k} + \frac{\beta_4/(\beta_2^2k)}{\beta_2s + \beta_3} \quad (22)$$

$$= \frac{d_0}{\beta_2k}\,s + \left(k\beta_3 + \frac{s\beta_4k}{\beta_1s + \beta_2}\right)^{-1}. \quad (23)$$

The decomposition in (22) is obtained by subtracting off the minimum of the real part of the second term in (21), and the same procedure is used on the inverse of that term to give (23). Equations (22) and (23) together with (20) give the realizations shown in Figs. 13 and 14 with the following expressions for the constants: $k_b = k\beta_2/d_0$:

$$c_1 = \frac{k\beta_2^2\beta_3}{\beta_4} \quad c_2 = \frac{k\beta_2^2}{\beta_1} \quad b_1 = \frac{k\beta_2^3}{\beta_4} \quad (24)$$

$$c_3 = k\beta_3 \quad c_4 = \frac{k\beta_4}{\beta_1} \quad b_2 = \frac{k\beta_4}{\beta_2}. \quad (25)$$

We can observe that

$$c_3 = c_1\frac{\beta_4}{\beta_2^2}, \qquad c_4 = c_2\frac{\beta_4}{\beta_2^2}, \qquad b_2 = b_1\left(\frac{\beta_4}{\beta_2^2}\right)^2.$$

Since $0 < \beta_4/\beta_2^2 \le 1$ we see that $c_3 \le c_1$, $c_4 \le c_2$ and $b_2 \le b_1$, so the realization of Fig. 14 is the more efficient one in the sense of having smaller parameter values.

It is useful to point out that the Brune procedure was relatively simple for the admittance (13), since the minimum real part occurred at zero or infinite frequency where the imaginary part was zero. Thus, the use of a "Brune cycle" involving transformers, or the alternative Bott–Duffin procedure, was not required.

E. A Strut Design Example

To illustrate the potential application of the inerter we consider the simple idealized problem of designing a suspension strut which has high static stiffness to applied loads F_s but which has well-damped time responses. We choose the following parameters: $m_s = 250$ kg, $m_u = 35$ kg, $k_t = 150$ kNm^{-1} and require that the strut behaves statically like a spring of stiffness $k_h = 120$ kNm^{-1} [27], [29]. We consider the set of system poles of the quarter-car model of (7), (8), which is equal to the set of zeros of

$$1 + Q(s)\,\frac{(m_u + m_s)s^2 + k_t}{m_ss(m_us^2 + k_t)}. \quad (26)$$

We consider the least damping ratio $\zeta_{\min}$ among all the system poles for a given $Q(s)$. We seek to maximize $\zeta_{\min}$ as a function of the admittance $Q(s)$ for various choices of admittance classes.

1) Design 1: SD Admittance With One Damper: We consider the case of $Q(s) = Y_1(s)$ as in (11) with $k = k_h$. The optimization of $\zeta_{\min}$ over T_1 and T_2 appears to be convex (see Fig. 15) with a maximum at $T_2 = 7.1465 \times 10^{-2}$, $T_1 = 0$ and

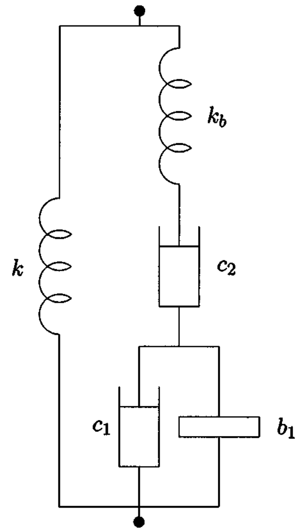

Fig. 13. First realization of the admittance (13).

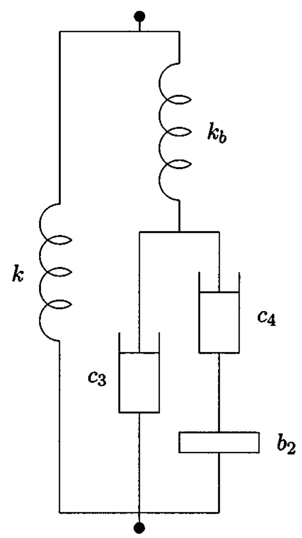

Fig. 14. Second realization of the admittance (13).

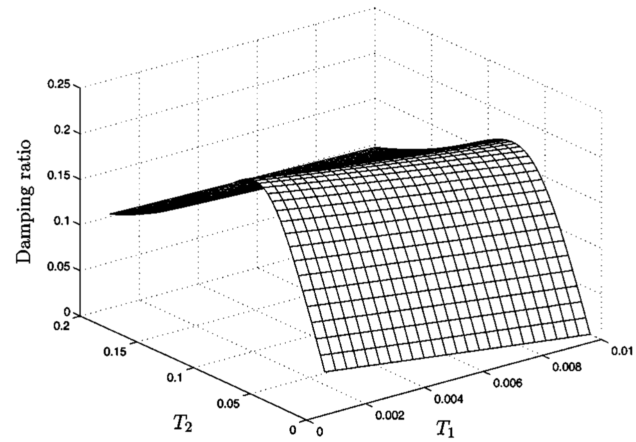

Fig. 15. Plot of damping ratio $\zeta_{\min}$ versus T_1 and T_2 in Design 1.

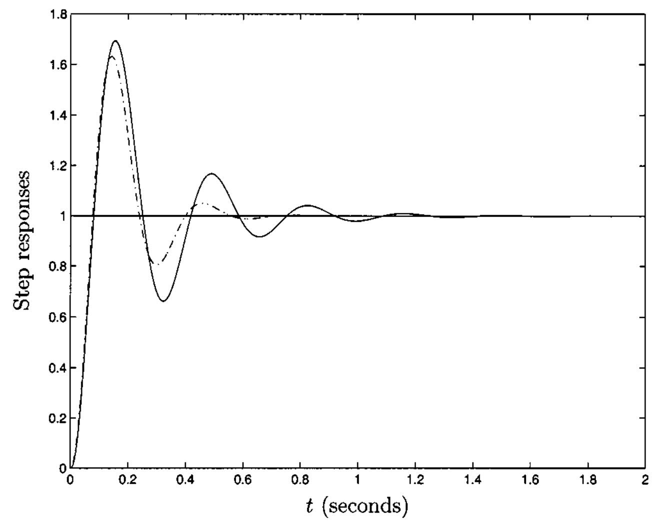

Fig. 16. Response of z_s to a unit step at z_r: Design 1 (–) and Design 3 (-·).

$\zeta_{\min} = 0.218$. This corresponds to the simplest suspension strut of a spring in parallel with a damper with admittance

$$Q(s) = \frac{sc_h + k_h}{s}$$

where the damper constant is given by $c_h = 8.58$ kNsm^{-1}. Fig. 16 shows the step response from z_r to z_s with rather light damping in evidence. This highlights one of the difficulties with conventional suspension struts which are very stiffly sprung.

2) Design 2: SD Admittance With Multiple Dampers: For the same optimization problem, but with $Q(s) = Y_2(s)$ as in (12) and $k = k_h$, direct searches using the Nelder–Mead simplex method led to no improvement on the value of $\zeta_{\min} = 0.218$ obtained in Design 1 with one damper. This situation appears to persist for a higher number of dampers as in Fig. 12. Further direct searches in the parameter space converged toward a set of pole-zero cancellations, and consequent reduction in McMillan degree in $Q(s)$, leaving only the solution obtained in Design 1. These claims are backed only by computational evidence, with a formal proof being lacking.

3) Design 3: Degree 3 Positive-Real Admittance: For the same optimization problem, but with the admittance $Q(s) = Y(s)$ as in equation (13), $k = k_h$ and no SD restriction, direct searches using the Nelder–Mead simplex method led to clear improvements. The following parameters

$$\begin{aligned} a_0 &= 1.8415 \times 10^{-3} \quad a_1 = 8.6105 \times 10^{-2} \\ d_0 &= 0 \quad d_1 = 4.0112 \times 10^{-3} \end{aligned} \tag{27}$$

gave a value of $\zeta_{\min} = 0.481$. The improved damping is demonstrated in Fig. 16 compared to the case of Design 1. The positive-real nature of $Y(s)$ is illustrated in the Bode plot of Fig. 17, which also clearly shows that this solution is employing phase advance. The latter fact proves that $Y(s)$ is not an SD admittance, as is also seen by $\beta_4 = 2.7847 \times 10^{-6} > 0$. The values for the constants in the realizations of Figs. 13 and 14 are given by $k_b = \infty$

$$\begin{aligned} c_1 &= 11.996 \text{ kNsm}^{-1} \quad c_2 = 55.091 \text{ kNsm}^{-1} \\ b_1 &= 269.10 \text{ kg} \quad c_3 = 9.8513 \text{ kNsm}^{-1} \\ c_4 &= 45.239 \text{ kNsm}^{-1} \quad b_2 = 181.46 \text{ kg.} \end{aligned}$$

These values appear to be within the bounds of practicality.

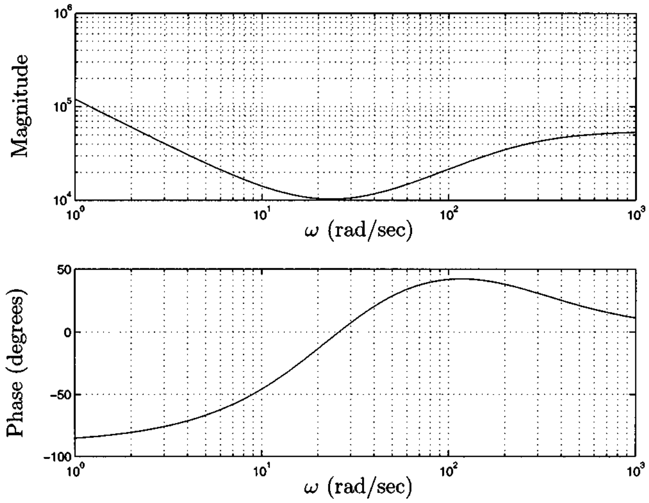

Fig. 17. Bode plot for the admittance $Y(s)$ of Design 3.

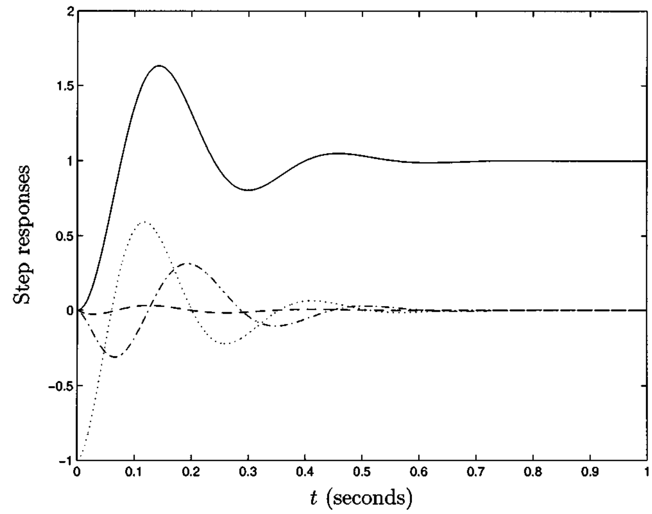

Fig. 18. Responses of Design 3 to a step input at z_r: z_s (–), $z_s - z_u$ (-·), $z_u - z_r$ (·), and deflection of damper c_4 (– –).

Fig. 18 shows the response of the sprung mass, suspension working space, tire deflection, and relative displacement of the damper c_4 (in the realization of Fig. 14) to a unit step road disturbance. Note that the inerter linear travel has a similar overall magnitude to the strut deflection due to the fact that the damper c_4 is quite stiff and has small travel.

F. Realizations Using Darlington Synthesis

The realizations shown in Figs. 13 and 14 both require the use of two dampers. It is interesting to ask if the admittance (13) may be realized using only one damper. An approach which will achieve this uses the method of Darlington [5], [11, Ch. IX.6]. In the electrical context the method allows any positive-real function to be realized as the driving-point impedance of a lossless two-port network terminated in a single resistance as shown in Fig. 19. Since there is no *a priori* estimate on the minimum number of inductors, capacitors (and indeed transformers) required for the realization of the lossless network, we will need to carry out the procedure to determine if the saving of one damper is offset by other increases of complexity, e.g., the need for more than one inerter or the use of levers.

For a reciprocal two-port network with impedance matrix

$$Z = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{12} & Z_{22} \end{pmatrix}$$

we can check that in Fig. 19

$$Z_1(s) = Z_{11} \frac{R^{-1}(Z_{11}Z_{22} - Z_{12}^2)/Z_{11} + 1}{R^{-1}Z_{22} + 1}. \tag{28}$$

Writing

$$Z_1 = \frac{m_1 + n_1}{m_2 + n_2} = \frac{n_1}{m_2} \frac{m_1/n_1 + 1}{n_2/m_2 + 1}$$

where m_1, m_2 are polynomials of even powers of s and n_1, n_2 are polynomials of odd powers of s, suggests the identification

$$Z_{11} = \frac{n_1}{m_2}, \quad Z_{22} = R\frac{n_2}{m_2} \quad Z_{12} = \sqrt{R}\,\frac{\sqrt{n_1 n_2 - m_1 m_2}}{m_2}. \tag{29}$$

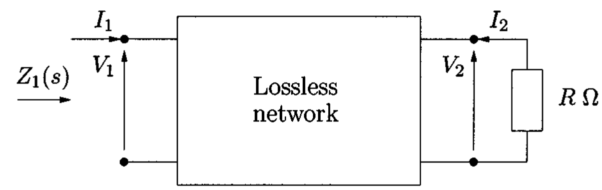

Fig. 19. Realization of a positive real impedance in Darlington form.

This corresponds to Case B in [11, Ch. IX.6]. Defining $Z_1(s) = Y(s)$ in (13) gives $m_1 = d_1 s^2$, $n_1 = s(d_0 s^2 + 1)$, $m_2 = k(a_0 s^2 + 1)$, and $n_2 = k a_1 s$, from which

$$n_1 n_2 - m_1 m_2 = k s^2(\beta_3 - \beta_1 s^2). \tag{30}$$

We now make the following two assumptions:

A1) $a_0 > 0$;
A2) $\min(\beta_1, \beta_3) > 0$.

The first of these involves no loss of generality since $a_0 = 0$ implies $d_0 = 0$, which means that (13) is an SD admittance requiring only one damper for its realization. The second avoids various special cases which can be dealt with in a rather simpler way than the general development which now follows.

Since (30) is not a perfect square it is necessary to multiply numerator and denominator in $Z_1(s)$ by the polynomial $P_0 = m_0 + n_0 = \sqrt{\beta_3} + \sqrt{\beta_1}\, s$ (determined by $m_0^2 - n_0^2 = \beta_3 - \beta_1 s^2$). This gives $Z_1(s) = (m_1' + n_1')/(m_2' + n_2')$ where

$$\begin{aligned}
m_1'(s) &= s^2\left(d_0\sqrt{\beta_1}\, s^2 + d_1\sqrt{\beta_3} + \sqrt{\beta_1}\right)\\
n_1'(s) &= s\left(\left(d_0\sqrt{\beta_3} + d_1\sqrt{\beta_1}\right)s^2 + \sqrt{\beta_3}\right)\\
m_2'(s) &= k\left(\left(a_0\sqrt{\beta_3} + a_1\sqrt{\beta_1}\right)s^2 + \sqrt{\beta_3}\right)\\
n_2'(s) &= ks\left(a_0\sqrt{\beta_1}\, s^2 + a_1\sqrt{\beta_3} + \sqrt{\beta_1}\right)
\end{aligned}$$

and $n_1' n_2' - m_1' m_2' = k s^2(\beta_3 - \beta_1 s^2)^2$. Then, using the correspondences in (29) with m_1 replaced by m_1', etc., we obtain the following expression for the impedance of the lossless two-port:

$$Z(s) = \frac{1}{m_2'}\begin{pmatrix} n_1' & \sqrt{Rk}\, s(\beta_3 - \beta_1 s^2) \\ \sqrt{Rk}\, s(\beta_3 - \beta_1 s^2) & R n_2' \end{pmatrix}.$$

We now write

$$Z(s) = sC_1 + \frac{s}{\left(a_0\sqrt{\beta_3} + a_1\sqrt{\beta_1}\right)s^2 + \sqrt{\beta_3}}C_2 \quad (31)$$

where the constant matrices C_1 and C_2 are given by

$$C_1 = \begin{pmatrix} \frac{d_0\sqrt{\beta_3}+d_1\sqrt{\beta_1}}{k\left(a_0\sqrt{\beta_3}+a_1\sqrt{\beta_1}\right)} & \frac{-\sqrt{R}\beta_1}{\sqrt{k}\left(a_0\sqrt{\beta_3}+a_1\sqrt{\beta_1}\right)} \\ \frac{-\sqrt{R}\beta_1}{\sqrt{k}\left(a_0\sqrt{\beta_3}+a_1\sqrt{\beta_1}\right)} & \frac{Ra_0\sqrt{\beta_1}}{a_0\sqrt{\beta_3}+a_1\sqrt{\beta_1}} \end{pmatrix}$$

$$C_2 = \frac{\beta_2+\sqrt{\beta_1\beta_3}}{a_0\sqrt{\beta_3}+a_1\sqrt{\beta_1}}\begin{pmatrix} \frac{\beta_3}{k} & \frac{\sqrt{R\beta_3}\,a_1}{\sqrt{k}} \\ \frac{\sqrt{R\beta_3}\,a_1}{\sqrt{k}} & Ra_1^2 \end{pmatrix}.$$

Following [11, Ch. VII.1] we note that any impedance matrix of the form $f(s)H$, where $f(s)$ is scalar and

$$H := \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix}$$

is nonnegative definite, can be realized in the form of the T-circuit of Fig. 20 with $X_i = \mu_i f(s)$ and

$$\mu_1 = h_{11} - \frac{h_{12}}{n} \quad \mu_2 = \frac{h_{22}}{n^2} - \frac{h_{12}}{n} \quad \mu_3 = \frac{h_{12}}{n}.$$

In order for $\mu_i \geq 0$, it is necessary and sufficient that $\operatorname{sign}(n) = \operatorname{sign}(h_{12})$ and

$$\frac{|h_{12}|}{h_{11}} \leq |n| \leq \frac{h_{22}}{|h_{12}|}.$$

We can now apply this realization procedure to each term in (31) and obtain the sum by taking a series connection of the two resulting two-ports (see Fig. 21).

We begin with C_2 and observe that $\det(C_2) = 0$, which fixes the choice of n and ensures that $\mu_1 = \mu_2 = 0$. Since both X_1 and X_2 then vanish from the T-circuit, this has the considerable advantage that only a single oscillator (inductor–capacitor or spring–inerter parallel connection) will be required, which is a significant economy. Moreover, we may dispense with a transformer by a choice of R which gives $n = 1$:

$$R = \frac{\beta_3}{ka_1^2} =: R_1. \quad (32)$$

We now set the first element of the second term in (31) equal to a parallel capacitor–inductor combination with impedance $(bs + k_1/s)^{-1}$ to obtain the parameters

$$b = \frac{k\left(a_0\sqrt{\beta_3}+a_1\sqrt{\beta_1}\right)^2}{\beta_3\left(\beta_2+\sqrt{\beta_1\beta_3}\right)} \quad (33)$$

$$k_1 = \frac{k\left(a_0\sqrt{\beta_3}+a_1\sqrt{\beta_1}\right)}{\sqrt{\beta_3}\left(\beta_2+\sqrt{\beta_1\beta_3}\right)}. \quad (34)$$

We denote the capacitance by b and the inductance by k_1^{-1} in anticipation of the mechanical analogy.

Turning to C_1, we note that a transformer will be required here with $n < 0$ and the range of transformer ratios given by

$$\frac{\beta_1\sqrt{\beta_3}}{a_1\left(d_0\sqrt{\beta_3}+d_1\sqrt{\beta_1}\right)} \leq |n| \leq \frac{a_0}{a_1}\sqrt{\frac{\beta_3}{\beta_1}}. \quad (35)$$

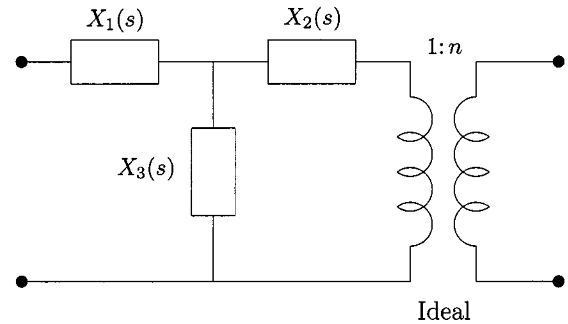

Fig. 20. T-circuit realization of elementary lossless two-ports.

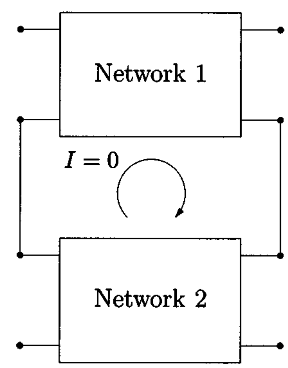

Fig. 21. Series connection of a pair of two-port networks.

Choosing $|n|$ at the lower (respectively, upper) limit sets X_1 (respectively, X_2) to be zero. It will be convenient to always select the lower limit. After suitable manipulation we can find expressions for the inductances $\mu_2 = k_2^{-1}$ and $\mu_3 = k_3^{-1}$ (whose inverses will be spring constants in the mechanical analogy)

$$k_2 = \frac{k\beta_1\sqrt{\beta_1}}{d_0\left(d_0\sqrt{\beta_3}+d_1\sqrt{\beta_1}\right)} \quad (36)$$

$$k_3 = k\frac{a_0\sqrt{\beta_3}+a_1\sqrt{\beta_1}}{d_0\sqrt{\beta_3}+d_1\sqrt{\beta_1}}. \quad (37)$$

We can now take the series connection of the two terms in (31) plus the terminating resistor to yield the circuit diagram shown in Fig. 22. A technical condition for the series connection in Fig. 21 to be valid is that no circulatory current can exist, which is satisfied in this case because of the presence of a transformer (see [11, Ch. VI.1], [13, pp. 325–326]).

It remains to deduce the mechanical analog of Fig. 22. The ideal transformer can be implemented as a simple lever with pivot point at the common node of the spring–inerter parallel combination. The central pivot automatically corresponds to an ideal transformer with negative turns ratio, which is what we require. Therefore, we obtain the mechanical realization of the admittance (13) with b, k_1, k_2, and k_3 given by (33), (34), (36), and (37)

$$c = \frac{ka_1^2}{\beta_3} \quad \text{and} \quad \frac{\lambda_2}{\lambda_1} = \frac{\beta_1\sqrt{\beta_3}}{a_1\left(d_0\sqrt{\beta_3}+d_1\sqrt{\beta_1}\right)} =: \rho. \quad (38)$$

It is possible to directly calculate the admittance of the mechanical one-port networks in Fig. 23 as a function of the parameters k_1, k_2, k_3, b, c, ρ, to be given by (39) at the bottom of the next page. Clearly, the McMillan degree of (39) is one higher

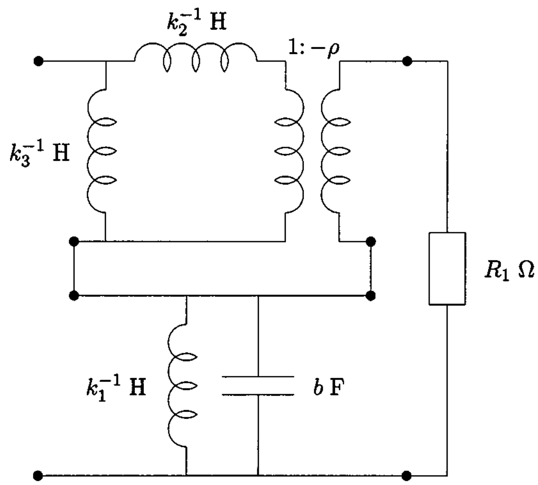

Fig. 22. Electrical circuit realization of the admittance (13).

than the admittance we started with in (13). Since there are four energy storage elements in Fig. 23 (three springs and one inerter), the extra degree is not unexpected from general circuit theory considerations. How then is equality with (13) to be explained? The answer is that there is an interdependence in the parameter values of k_1, k_2, k_3, b, c, ρ, as defined through (33), (34), and (36)–(38) which is sufficient to ensure a pole-zero cancellation in (39). In the case when $d_0 = 0$, which makes $k_2 = \infty$, this interdependence is expressed by

$$(k_1\rho^3 + (2k_1 - k_3)\rho^2 + (k_1 - 2k_3)\rho - k_3)\rho c^2 + bk_3^2 = 0.$$

(The general relationship is significantly more complicated). It is evident that the mechanical network in Fig. 23 parameterizes a class of admittances which is strictly larger than those in (13) if the parameter values of k_1, k_2, k_3, b, c, and ρ are allowed to vary independently.

It is interesting to make any possible comparisons between parameter values required in Fig. 23 and those for the realization in Fig. 14 for the admittance (13). In fact, it is possible to show

$$b_2 < b \tag{40}$$
$$c_3 \le c \tag{41}$$
$$k \le k_1 \tag{42}$$
$$k_b \ge k_2. \tag{43}$$

To show (40), we note that $b_2 < k\beta_2 \le ka_0$ while

$$b = \frac{ka_0\left(a_0\beta_3 + 2a_1\sqrt{\beta_1\beta_3} + a_1^2\beta_1/a_0\right)}{(a_0 - d_0)\beta_3 + (a_1 - d_1)\sqrt{\beta_1\beta_3}} > ka_0.$$

For (41), note that $c_3 = k(a_1 - d_1)$ and $c = ka_1^2/(a_1 - d_1)$; (42) follows from:

$$k_1 = \frac{k\left(a_0\sqrt{\beta_3} + a_1\sqrt{\beta_1}\right)}{(a_0 - d_0)\sqrt{\beta_3} + (a_1 - d_1)\sqrt{\beta_1}}$$

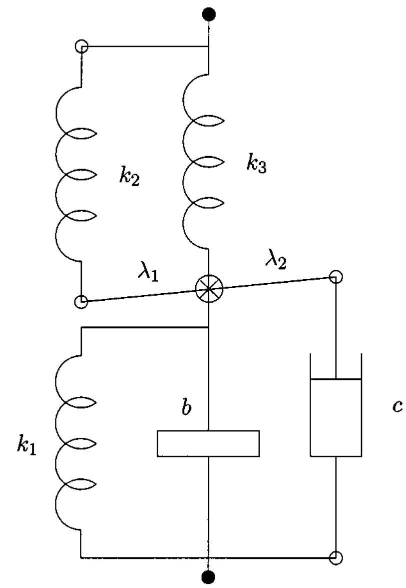

Fig. 23. Third realization of the admittance (13).

and (43) follows from:

$$\begin{aligned}(k_b - k_2)&\frac{d_0\left(d_0\sqrt{\beta_3} + d_1\sqrt{\beta_1}\right)}{k}\\ &= \beta_2\left(d_0\sqrt{\beta_3} + d_1\sqrt{\beta_1}\right) - \beta_1\sqrt{\beta_1}\\ &\ge d_0\left(\beta_2\sqrt{\beta_3} + \beta_3\sqrt{\beta_1}\right) \ge 0.\end{aligned}$$

It appears difficult to give any useful *a priori* estimates for the lever ratio ρ.

We now return to the suspension strut design of Section IV-E-3. For the parameter values given in (27), the realization of Fig. 23 gives the following values for the constants using equations (33), (34), and (36)–(38)

$$\begin{aligned}b &= 323.62\ \text{kg} \quad k_1 = 121.74\ \text{kNm}^{-1} \quad k_2 = \infty\\ k_3 &= 8384.7\ \text{kNm}^{-1} \quad c = 10.837\ \text{kNsm}^{-1} \quad \rho = 2.2546.\end{aligned}$$

These values appear to be within the bounds of practicality.

After carrying out the Darlington procedure for the realization of the admittance (13), we note that the saving of one damper from the realizations of Figs. 13 and 14 has been offset by the need for a lever. The extra spring and increased value of inerter constant implied in (40) are perhaps less significant differences.

V. Simulated Mass

The previous two applications of the inerter exploited one of its principal advantages over the mass element, namely that neither of its terminals need to be grounded. There is also the possibility that the inerter could be used to replace a mass

$$\frac{s^3(k_2 + k_3)cb\rho^2 + s^2k_2k_3b + s(\rho^2k_1(k_2 + k_3) + k_2k_3)c + k_1k_2k_3}{s(s^3cb\rho^2 + s^2bk_2 + sc(\rho^2(k_1 + k_3) + (1 + \rho)^2k_2) + k_2(k_1 + k_3))}. \tag{39}$$

element with one of its terminals then being connected to ground. This is illustrated in Fig. 24(a) and (b), which are in principle equivalent dynamically with respect to displacement disturbances z.

Fig. 24(b) may be a useful alternative to Fig. 24(a) in a situation where it is desired to test a spring–damper support or absorber before final installation, and where it is impractical to test it on a real mass element, e.g., where the mass M is very large.

By contrast, it should be pointed out that, even in the context of mechanical network synthesis, Fig. 24(b) may not be a physically feasible alternative to Fig. 24(a) in situations where it is impossible to connect one terminal of the inerter to ground, e.g., for a vibration absorber mounted on a bridge.

VI. Conclusion

This paper has introduced the concept of the ideal inerter, which is a two-terminal mechanical element with the defining property that the relative acceleration between the two terminals is proportional to the force applied on the terminals. There is no restriction that either terminal be grounded, i.e., connected to a fixed point in an inertial frame. The element may be assumed to have small or negligible mass. The ideal inerter plays the role of the true network dual of the (ideal) mechanical spring.

It was shown that the inerter is capable of simple realization. One approach is to take a plunger sliding in a cylinder driving a flywheel through a rack, pinion and gears. Such a realization satisfies the property that no part of the device need be attached to ground, and that it has a finite linear travel which is specifiable. The mass of the device may be kept small relative to the inertance (constant of proportionality) by employing a sufficiently large gear ratio. Such a realization may be viewed as approximating its mathematical ideal in the same way that real springs, dampers, capacitors, etc. approximate their mathematical ideals.

The inerter completes the triple of basic mechanical network elements in a way that is advantageous for network synthesis. The properties that neither terminal need be grounded and the device mass may be small compared to the inertance are crucial for this purpose. It allows classical electrical circuit synthesis to be exploited directly to synthesize any one-port (real-rational) positive-real impedance as a finite network comprising springs, dampers, and inerters. The use of the inerter for synthesis does not prevent mechanical networks containing mass elements from being *analyzed* in the usual way as the analogs of grounded capacitors. Moreover, as well as the possibility that in some situations it is advantageous that one terminal of the mass element is the ground, there is also the possibility that the inerter may have benefits to *simulate* a mass element with one of its terminals being connected to ground.

A vibration absorption problem was considered as a possible application of the inerter. Rather than mounting a tuned spring–mass system on the machine that is to be protected from oscillation (conventional approach), a black-box mechanical admittance was designed to support the machine with a blocking zero on the imaginary axis at the appropriate frequency. The resulting mechanical network consisted of a parallel spring-damper in series with a parallel spring inerter. This arrangement avoids any associated problems of attaching the spring-mass to the machine, such as the need for an undesirably large mass to limit its travel.

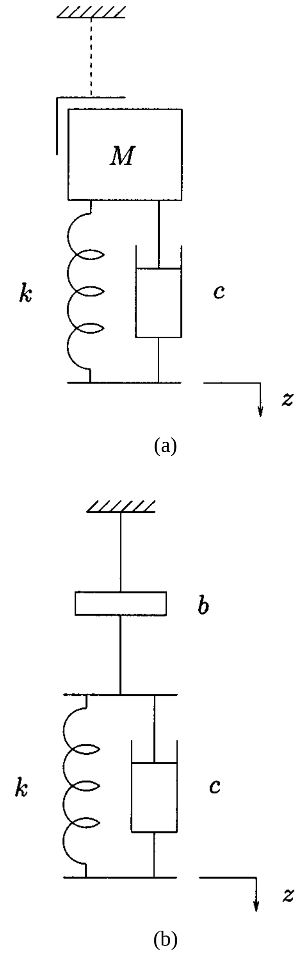

Fig. 24. Spring–damper supporting (a) a mass element and (b) a grounded inerter acting as a simulated mass.

A vehicle suspension strut design problem was considered as another possible application of the inerter. It was pointed out that conventional struts comprising only springs and dampers have severely restricted admittance functions, namely their poles and zeros all lie on the negativ real axis and the poles and zeros alternate, so that the admittance function always has a lagging frequency response. The problem of designing a suspension strut with very high static spring stiffness was considered. It was seen that conventional spring and damper arrangements always resulted in very oscillatory behavior, but the use of inerters can reduce the oscillation. In studying this problem, a general positive real admittance was considered consisting of two zeros and three poles. The realization procedure of Brune was applied to give two circuit realizations of the admittance, each of which consisted of two springs, two dampers and one inerter. The resulting parameter values for the strut design appear within the bounds of practicality. As an alternative, the realization procedure of Darlington was used to finding a realization consisting of one damper, one inerter, three springs and a lever.

Acknowledgment

The author would like to thank G. W. Walker for many stimulating conversations on vehicle suspensions and mechanical networks. It was within the context of these conversations that the question was asked on the possible existence of a physical device of the type defined here, which was later found to be realizable, and coined "the inerter."

References

[1] B. D. O. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis*. Upper Saddle River, NJ: Prentice-Hall, 1973.
[2] V. Belevitch, *Classical Network Theory*. San Francisco, CA: Holden-Day, 1968.
[3] R. Bott and R. J. Duffin, "Impedance synthesis without use of transformers," *J. Appl. Phys.*, vol. 20, p. 816, 1949.
[4] O. Brune, "Synthesis of a finite two-terminal network whose driving-point impedance is a prescribed function of frequency," *J. Math. Phys.*, vol. 10, pp. 191–236, 1931.
[5] S. Darlington, "Synthesis of reactance 4-poles which produce prescribed insertion loss characteristics," *J. Math. Phys.*, vol. 10, pp. 257–353, 1931.
[6] ——, "A history of network synthesis and filter theory for circuits composed of resistors, inductors, and capacitors," *IEEE Trans. Circuits Syst.I*, vol. 46, pp. 4–13, Jan. 1999.
[7] M. Darrieus, "Les modèles mécaniques en électrotechnique leur application aux problèmes de stabilité," *Bull. Soc. Fra. Elect.*, vol. 96, pp. 794–809, 1929.
[8] J. P. Den Hartog, *Mechanical Vibrations*. New York: Dover, 1985.
[9] F. A. Firestone, "A new analogy between mechanical and electrical systems," *J. Acoust. Soc. Amer.*, vol. 4, pp. 249–267, 1933.
[10] F. R. Gantmacher, *The Theory of Matrices*. New York: Chelsea, 1980, vol. II.
[11] E. A. Guillemin, *Synthesis of Passive Networks*. New York: Wiley, 1957.
[12] W. Hähnle, "Die Darstellung elektromechanischer Gebilde durch rein elektrische Schaltbilder," *Wiss. Veröff. Siemens-Konz.*, vol. 11, pp. 1–23, 1932.
[13] W. W. Harman and D. W. Lytle, *Electrical and Mechanical Networks*: McGraw-Hill, 1962.
[14] J. K. Hedrick and T. Butsuen, "Invariant properties of automotive suspensions," *Proc. Inst. Mech. Eng D: J. Automobile Eng.*, vol. 204, pp. 21–27, 1990.
[15] E. L. Hixson, "Mechanical impedance," in *Shock and Vibration Handbook*, 2nd ed, C. M. Harris and C. E. Crede, Eds. New York: McGraw-Hill, 1976, ch. 10.
[16] D. C. Karnopp and R. C. Rosenberg, *Analysis and Simulation of Multiport Systems: The Bond Graph Approach to Physical System Dynamics*. Cambridge, MA: MIT Press, 1968.
[17] D. C. Karnopp, D. L. Margolis, and R. C. Rosenberg, *System Dynamics: Modeling and Simulation of Mechatronic Systems*. New York: Wiley, 2000.
[18] W. J. Karplus and W. W. Soroka, *Analog Methods*. New York: McGraw-Hill, 1959.
[19] H. E. Koenig and W. A. Blackwell, *Electromechanical System Theory*. New York: McGraw-Hill, 1961.
[20] A. G. J. MacFarlane, *Dynamical System Models*. Edinburgh, U.K.: Chambers Harrap, 1970.
[21] R. W. Newcomb, *Linear Multiport Synthesis*. New York: McGraw-Hill, 1966.
[22] H. F. Olson, *Dynamical Analogies*. New York: Van Nostrand, 1943.
[23] H. M. Paynter, *Analysis and Design of Engineering Systems*. Cambridge, MA: MIT Press, 1961.
[24] J. C. Schönfeld, "Analogy of hydraulic, mechanical, acoustic and electrical systems," *Appl. Sci. Res., B*, vol. 3, pp. 417–450, 1954.
[25] R. S. Sharp and S. A. Hassan, "The relative performance capabilities of passive, active and semi-active car suspension systems," in *Proc. Inst. Mech. Engs.*, vol. 200, 1986, pp. 219–228.
[26] J. L. Shearer, A. T. Murphy, and H. H. Richardson, *Introduction to System Dynamics*. Reading, MA: Addison-Wesley, 1967.
[27] M. C. Smith, "Achievable dynamic response for automotive active suspension," *Veh. Syst. Dyna.*, vol. 24, pp. 1–33, 1995.
[28] M. C. Smith and G. W. Walker, "A mechanical network approach to performance capabilities of passive suspensions," in *Proceedings of the Workshop on Modeling and Control of Mechanical Systems*. London, U.K.: Imperial College Press, June 17–20, 1997, pp. 103–117.
[29] M. C. Smith and G. W. Walker, "Performance limitations and constraints for active and passive suspensions: A mechanical multi-port approach," *Veh. Syst. Dyna.*, vol. 33, pp. 137–168, 2000.
[30] H. M. Trent, "Isomorphisms between oriented linear graphs and lumped physical systems," *J. Acoust. Soc. Amer.*, vol. 27, pp. 500–527, 1955.
[31] P. E. Wellstead, *Introduction to Physical System Modeling*. New York: Academic, 1979.

Malcolm Smith (M'90–SM'00–F'02) received the B.A. degree in mathematics, the M.Phil. degree in control engineering and operational research, and the Ph.D. degree in control engineering from the University of Cambridge, Cambridge, U.K., in 1978, 1979, and 1982, respectively.

He was subsequently a Research Fellow at the German Aerospace Center, DLR, Oberpfaffenhofen, Germany, a Visiting Assistant Professor and Research Fellow with the Department of Electrical Engineering at McGill University, Montreal, Canada, and an Assistant Professor with the Department of Electrical Engineering at the Ohio State University, Columbus, OH. In 1990, he returned to Cambridge University as a Lecturer in the Department of Engineering, and became a Reader in 1997. He is a Fellow of Gonville and Caius College, Cambridge, U.K. His research interests include control system design, frequency response methods, H-infinity optimization, nonlinear systems, active suspension, and mechanical systems.

Dr. Smith was a corecipient of the George Axelby Outstanding Paper Award in the IEEE TRANSACTIONS ON AUTOMATIC CONTROL in 1992 and 1999, both times for joint work with T. T. Georgiou.