The Missing Mechanical Circuit Element

Michael Z.Q. Chen, Christos Papageorgiou, Frank Scheibe, Fu-Cheng Wang, and Malcolm C. Smith

Abstract

In 2008, two articles in Autosport revealed details of a new mechanical suspension component with the name “J-damper” which had entered Formula One Racing and which was delivering significant performance gains in handling and grip. From its first mention in the 2007 Formula One “spy scandal” there was much speculation about what the J-damper actually was. The Autosport articles revealed that the J-damper was in fact an “inerter” and that its origin lay in academic work on mechanical and electrical circuits at Cambridge University. This article aims to provide an overview of the background and origin of the inerter, its application, and its intimate connection with the classical theory of network synthesis.
1. Introduction

The standard analogies between mechanical and electrical networks are universally familiar to students and engineers alike. The basic modelling elements have the following correspondences:

- spring ↔ inductor
- damper ↔ resistor
- mass ↔ capacitor,

where force relates to current and velocity to voltage. It is known that the correspondence is perfect in the case of the spring and damper. A fact which is also known, but frequently glossed over, is that there is a restriction in the case of the mass. All the above elements except the mass have two “terminals” (for a mechanical element the terminals are the attachment points which should be freely and independently movable in space). In contrast, the mass element has only one such terminal—the centre of mass. It turns out that the mass element is analogous to a grounded electrical capacitor (see Sidebar I). The above correspondence is so familiar that one does not think to question it. However, a careful examination of the force-current analogy with the inerter replacing the mass element. The admittance \( Y(s) \) is the ratio of through to across quantities, where \( s \) is the standard Laplace transform variable. For mechanical networks in rotational form the through and across variables are torque and angular velocity, respectively. For further background on network analogies see [23], [35], and [38].

**Figure 1.** Electrical and mechanical circuit symbols and correspondences. In the force-current analogy forces substitute for currents and velocities substitute for voltages. The admittance \( Y(s) \) maps velocity and voltage into force and current, respectively. (The symbol \( s \) is the standard Laplace transform variable.)
As an ideal modelling element, the inerter is defined to be a two-terminal mechanical device such that the applied force at the terminals is proportional to the relative acceleration between them.

of the classical theory of electrical networks suggests otherwise. The famous result of Bott and Duffin [3] says that an arbitrary passive driving-point impedance can be realized as a two-terminal network comprising resistors, capacitors and inductors only. Since the mapping to mechanical circuits is power-preserving it is natural to expect that arbitrary passive mechanical impedances can be similarly realized. But there is a snag. A circuit in which neither terminal of a capacitor is grounded will not have a mechanical analogue. In applications where both mechanical terminals are movable (such as a vehicle suspension system) the restriction is a very real one.

To bypass the snag a new mechanical modelling element was proposed by Smith [38]. The element has two terminals, and has the property that the applied force at the terminals is proportional to the relative acceleration between them. It was shown that such devices can be built in a relatively simple manner [37], [38]. A new word “inerter” was coined to describe such a device. As well as offering new possibilities for “passive mechanical control” in a variety of applications, the inerter brought out strong connections with the classical theory of electrical circuit synthesis, reviving old questions and suggesting new ones.

Since the birth of the inerter in the Engineering Department at Cambridge University a number of applications have been proposed and investigated. Alongside the successful application in Formula One racing (see Sidebar II) the general applicability to vibration absorption and automotive suspensions has been considered [29], [38], [40]. The use of the inerter in mechanical steering compensators of high-performance motorcycles was studied in [14], [15]; by replacing the conventional steering damper with a serial inerter-damper layout, it was shown that two significant instabilities, “wobble” and “weave”, can be stabilized simultaneously. Further research saw the inerter proposed for train suspension systems [44], [46], in which the inerter was located in both the body-bogie and bogie-wheel connections. Recently, the inerter has been studied for building suspension control [43], where three building models being used to analyse the suspension performance. In all cases, the introduction of the inerter device has been shown to offer performance advantages over conventional passive solutions.

This article describes the background to the inerter, the connections with classical electrical circuit theory, and its applications. The rest of this article is organised as follows. Section II presents the physical constructions of the inerter. Section III reviews passive network synthesis, considers the suspension synthesis solution of restricted complexity, and presents a new test for positive-realness. Section IV presents positive-real synthesis using matrix inequalities and the analytical solutions for optimal ride comfort and tyre grip. In Section V the development of a simulation-based methodology is presented for the analysis and optimal design of nonlinear passive vehicle suspensions. Section VI presents a behavioural approach to play in mechanical networks. Conclusions are given in Section VII.

2. The Inerter and its Physical Embodiments

Let us focus attention first on the five familiar two-terminal modelling elements: resistor, capacitor, inductor, spring, and damper. Each is an ideal modelling element, with a precise mathematical definition. At the same time, each is a model for physical devices whose behaviour is an approximation to the ideal. The same is true for the inerter.

As an ideal modelling element, the inerter is defined to be a two-terminal mechanical device such that the applied force at the terminals is proportional to the relative acceleration between them. The constant of proportionality is called the inertance and has the units of kilograms. For this to be a useful definition, realistic embodiments are needed. The meaning of “realistic” was elaborated in [38]. It was argued that the inerter device should have a small mass relative to the inertance which should be adjustable independently of the mass. Also, the device should function properly in any spatial orientation, it should support adequate linear travel and it should have reasonable overall dimensions. Inerter with these features can be mechanically realized in various ways. In [38], a rack-and-pinion inerter (see Fig. 3(a)) was proposed using a flywheel that is driven by a rack and pinion, and gears. Other methods of construction are described in [37], e.g. using hydraulics or screw mechanisms. Fig. 3(b) shows a schematic of a ball-screw inerter and an example of such a device is pictured in Fig. 4. For such devices the value of the inertance which is easy to compute [37], [38]. In general, if the device gives rise to a flywheel rotation of radians per meter of relative displacement between the terminals, then the inertance of the device is given by where is the flywheel’s moment of inertia.
Like other modelling elements, the deviation of inerter embodiments from ideal behaviour should be kept in mind. Typical effects which have been observed and quantified include backlash, friction and elastic effects [20], [26], [27], [28], [45]. Backlash (mechanical play) in a physical inerter is a particularly interesting issue, theoretically and practically, which is discussed in Section VI.

### 3. Passive Network Synthesis

The literature on passive electrical network synthesis is both rich and vast. Excellent introductions to the field can be found in [1], [2], [17], [24], [42]. The concept of passivity can be translated over directly to mechanical networks as follows. Suppose that \((F, v)\) represents the force-velocity pair associated with a two-terminal mechanical network, then passivity requires:

\[
\int_{-\infty}^{T} F(t)v(t)dt \geq 0
\]

for all admissible time functions \(F(t), v(t)\) and all \(T\). If \(Z(s)\) is the real rational impedance or admittance function of a linear time-invariant two-terminal network, it is well-known that the network is passive if and only if \(Z(s)\) is positive-real [1], [24]. Let \(Z(s)\) be a real-rational function. Then \(Z(s)\) is defined to be positive-real if \(\text{Re}[Z(s)] \geq 0\) in the open right half plane (ORHP), i.e. for all \(s\) with \(\text{Re}[s] > 0\). The following is a well-known equivalent condition for positive-realness.
Theorem 1: [1], [24]: \( Z(1) \) is positive-real if and only if
1. \( Z(1) \) is analytic in \( \Re[s] > 0 \); and
2. \( \Re[Z(j\omega)] \geq 0 \) for all \( \omega \) with \( j\omega \) not a pole of \( Z(s) \);
3. poles on the imaginary axis and infinity are simple and have non-negative residues.

An alternative necessary and sufficient condition for positive-realness is as follows.

Theorem 2: [48], [49]: Let \( Z(s) = \frac{p(s)}{q(s)} \), where \( p(s) \) and \( q(s) \) are coprime polynomials. Then \( Z(s) \) is positive-real if and only if
1. \( p(s) + q(s) \) is Hurwitz;
2. \( \Re[Z(j\omega)] \geq 0 \) for all \( \omega \) with \( j\omega \) not a pole of \( Z(s) \).

In [3] Bott and Duffin showed that any rational positive-real function can be realized as the driving-point impedance of a two-terminal network comprising resistors, inductors, and capacitors only. Making use of the force-current analogy (see Sidebar I) and the new modelling element (inerter) it can be seen that, given any positive-real function \( Z(s) \), there exists a passive two-terminal mechanical network whose impedance equals \( Z(s) \), which consists of a finite interconnection of springs, dampers, and inerters. The ability to synthesise the most general positive-real impedance allows the designer to achieve the optimal performance among passive mechanical networks. Fig. 5 shows a specific mechanical network together with a physical realization constructed at Cambridge University Engineering Department.

Efficiency of realization, as defined by the number of elements used, is much more important for mechanical networks than electrical networks. In this section, we consider the class of realizations in which the number of dampers and inerters is restricted to one in each case while allowing an arbitrary number of springs (which is the easiest element to realize practically). Some examples of this class have been given in Figs. 10 and 12 (Section IV). This problem is analogous to restricting the number of resistors and capacitors, but not inductors, in electrical circuit synthesis [10]. Such questions involving restrictions on both resistive and one type of reactive element have never been considered. This contrasts with the problems of minimal resistive and minimal reactive synthesis which have well-known solutions when transformers are allowed ([13], [50], see also [1]). In our problem, we impose the condition that no transformers

Figure 3. Schematics of two embodiments of the inerter. (a) Rack and pinion inerter, (b) ballscrew inerter.

Figure 4. Ballscrew inerter made at Cambridge University Engineering Department; Mass = 1 kg, Inertance (adjustable) = 60-240 kg. (a) Complete with outer case, (b) ballscrew, nut and flywheel, (c) flywheel removed, (d) thrust bearing.
are employed, due to the fact that large lever ratios can give rise to practical problems. Such a case can occur if there is a specification on available “travel” between two terminals of a network, as in a car suspension. A large lever ratio may necessitate a large travel between internal nodes of a network, which then conflicts with packaging requirements.

We show that the problem considered here is closely related to the problem of one-element-kind multi-port synthesis. We then review the definition of paramountcy and its connection to transformerless synthesis. Five circuit realizations are then presented to cover the general class under consideration.

We consider a mechanical one-port network \( Q \) consisting of an arbitrary number of springs, one damper and one inerter. We can arrange the network in the form of Fig. 6 where \( X \) is a three-port network containing all the springs. The impedance matrix of \( X \) defined by

\[
\begin{bmatrix}
\dot{v}_1 \\
\dot{v}_2 \\
\dot{v}_3
\end{bmatrix} = s
\begin{bmatrix}
L_1 & L_4 & L_5 \\
L_4 & L_2 & L_6 \\
L_5 & L_6 & L_3
\end{bmatrix}
\begin{bmatrix}
\dot{F}_1 \\
\dot{F}_2 \\
\dot{F}_3
\end{bmatrix} = : sL \begin{bmatrix}
\dot{F}_1 \\
\dot{F}_2 \\
\dot{F}_3
\end{bmatrix},
\]

where \( L \) is a non-negative definite matrix and \( \dot{\cdot} \) denotes Laplace transform. And the admittance of \( Q \) is

\[
\frac{\dot{F}_1}{v_1} = \frac{\alpha_3 s^3 + \alpha_2 s^2 + \alpha_1 s + \alpha_0}{\beta_4 s^4 + \beta_3 s^3 + \beta_2 s^2 + \beta_1 s}, \tag{1}
\]

where \( \alpha_3 = bc(L_2L_3 - L_5^2), \quad \alpha_2 = bL_3, \quad \alpha_1 = cL_2, \quad \alpha_0 = 1, \quad \beta_4 = bc \det(L), \quad \beta_3 = b(L_1L_4 - L_3^2), \quad \beta_2 = c(L_1L_2 - L_6^2) \) and \( \beta_1 = L_1. \)

The admittance (1) effectively has only six parameters which can be adjusted among the seven coefficients. To see this note that \( b \) and \( c \) can be set to be equal to 1 and the following scalings carried out: \( L_1 \rightarrow R_1, cL_2 \rightarrow R_2, bL_3 \rightarrow R_3, \sqrt{c}L_4 \rightarrow R_4, \sqrt{b}L_5 \rightarrow R_5, \sqrt{bc}L_6 \rightarrow R_6 \), to leave (1) invariant. The resulting admittance is \( Y(s) = \frac{(R_3R_5 - R_6^2)s^3 + R_5s^2 + R_5s + 1}{s(\det Rs^3 + (R_3R_5 - R_6^2)s^2 + (R_5R_2 - R_4^2)s + R_4)} \tag{2} \)

and

\[
R = \begin{bmatrix}
R_1 & R_4 & R_5 \\
R_4 & R_2 & R_6 \\
R_5 & R_6 & R_3
\end{bmatrix} = T \begin{bmatrix}
L_1 & L_4 & L_5 \\
L_4 & L_2 & L_6 \\
L_5 & L_6 & L_3
\end{bmatrix} T,
\]

where

\[
T = \begin{bmatrix}
1 & 0 & 0 \\
0 & \sqrt{c} & 0 \\
0 & 0 & \sqrt{b}
\end{bmatrix}
\]

and \( R \) is non-negative definite.

We will now consider the conditions on \( L \) or \( R \) that will ensure that \( X \) corresponds to a network of springs only (and no transformers). To this end we introduce the following definition.

**Definition 1**: A matrix is defined to be paramount if its principal minors, of all orders, are greater than or equal to the absolute value of any minor built from the same rows \([6],[36]\).

It has been shown that paramountcy is a necessary condition for the realisability of an \( n \)-port resistive
network without transformers \cite{6}, \cite{36}. In general, paramountcy is not a sufficient condition for the realisability of a transformerless resistive network and a counter-example for $n = 4$ was given in \cite{7}, \cite{47}.

In \cite[pp. 166–168]{41}, it was proven that paramountcy is necessary and sufficient for the realisability of a resistive network without transformers with order less than or equal to three ($n \leq 3$). The construction of \cite{41} for the $n = 3$ case makes use of the network containing six resistors with judicious re-labelling of terminals and changes of polarity.

We now state a theorem from \cite{8}, \cite{9}, \cite{12} which provides specific realizations for the $Y(s)$ in the form of Fig. 6 for any $X$ that contains springs only and no transformers. The realizations are more efficient than would be obtained by directly using the construction of Tellegen in that only four springs are needed. This is due to the fact that Theorem 3 exploits the additional freedom in the parameters $b$ and $c$ to realize the admittance (2). Alternative realizations can also be found which are of similar complexity (see \cite{8}).

**Theorem 3**: \cite{8}, \cite{9}, \cite{12} Given $Y(s)$ in the form of Fig. 6 where $X$ contains only springs. Then $Y(s)$ can be realized with one damper, one inerter, and at most four springs in the form of Fig. 7(a)–7(e).

If we take a closer look at Eq. (2), it is a bi-cubic function multiplied by $1/s$. It appears difficult to determine necessary and sufficient condition for positive-realness of this class using existing results (Theorems 1 and 2). The convenient test provided by Theorem 2 is then no longer applicable and detailed checking of the residue conditions in Theorem 1 is still needed. This motivated the search for the improved test of Theorem 4.

**Theorem 4**: \cite{8}, \cite{11} Let $Z(s) = p(s)/q(s)$, where $p(s)$ and $q(s)$ have no common roots in the ORHP. Then $Z(s)$ is positive-real if and only if

1. $p(s) + q(s)$ has no roots in the ORHP;
2. $\text{Re}[Z(j\omega)] \geq 0$ for all $\omega$ with $j\omega$ not a pole of $Z(s)$.

When $p(s)$ and $q(s)$ are coprime, the “only if” implication is stronger in Theorem 2 than Theorem 4 while the reverse is the case for the “if” implication. The latter fact means that Theorem 4 is more powerful for testing the positive-realness of a given function. Although Theorem 4 appears only subtly different from Theorem 2 it gives a significant advantage, as seen in testing some classes of low-order positive-real functions \cite{8}, \cite{11}.

### 4. Vehicle Suspension

In general, a good suspension should provide a comfortable ride and good handling for a reasonable range of suspension deflections. The specific criteria used depend on the purpose of the vehicle. From a system design point of view, there are two main categories of disturbances on a vehicle, namely road and load disturbances (the latter being a simple approximation to
driver inputs in elementary vehicle models). Standard spectra are available to model stochastic road profile inputs. Load disturbances can be used to model forces induced by driver inputs such as accelerating, braking and cornering. In this way, suspension design can be thought of as a problem of disturbance rejection to selected performance outputs (e.g., vertical body acceleration, body pitch deflection, tyre deflection and suspension travel).

Passive suspensions contain elements such as springs, dampers, inerties and possibly levers, which can only store or dissipate energy, i.e. there is no energy source in the system. They therefore provide a simpler and cheaper means of suspension design and construction at the expense of performance limitations than active suspensions (with energy sources). Generally a suspension needs to be “soft” to insulate against road disturbances and “hard” to insulate against load disturbances. It is well-known that these objectives cannot be independently achieved with a passive suspension [21], [39]. However, the use of inerties in addition to springs and dampers can alleviate the necessary compromises between these two goals [29], [40].

In the next section, we show how suspension networks can be designed using a linear matrix inequality (LMI) approach (Section A). We also present some results on global optima which can be derived as a function of the quarter-car model parameters for some specific networks (Section B).

A. Design of Optimal Passive Suspension Networks

We summarize the approach of [29] where the suspension design problem was formulated as an optimal control problem over positive real admittances. The solution of the optimization problem made use of matrix inequalities and required the application of a local, iterative scheme due to the non-convexity of the problem. Even so, the design method was able to come up with new network topologies involving inerties that resulted in considerable improvement in the individual performance measures. It was also possible to formulate and solve multi-objective optimization problems.

1) The quarter-car model: The quarter-car model presented in Fig. 8 is the simplest model to consider for suspension design. It consists 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 by means of the positive-real admittance function \( Y(s) \) which relates the suspension force to the strut velocity. In this section we will assume further that \( Y(s) = K(s) + K_s/s \), where \( K(s) \) is positive-real and has no pole at \( s = 0 \) and \( K_s \) is fixed at the desired static stiffness. Here we fix the parameters of the quarter-car model as: \( m_s = 250 \text{ kg}, m_u = 35 \text{ kg}, \) and \( k_t = 150 \text{ kN/m} \).

2) The control synthesis paradigm: In order to synthesise admittances over the whole class of positive-real functions, we use a control synthesis paradigm along with a state-space characterisation of positive-realness. The search for positive-real admittances is formulated as a search for positive-real “controllers” \( K(s) \) as shown in Fig. 9 where \( w \) represents the exogenous disturbances (e.g. \( z_r \) and \( F_s \)) and \( z \) represents outputs to be controlled, e.g. sprung mass acceleration, tyre force, etc. The characterisation of positive-realness of the controller is achieved with the following result.
Lemma 2 (Positive real lemma [4]): Given that,
\[
K(s) = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix} = C_k (sI - A_k)^{-1} B_k + D_k,
\]
then \( K(s) \) is positive-real if and only if there exists \( P_k > 0 \) that satisfies the Linear Matrix Inequality (LMI)
\[
\begin{bmatrix} A_k^T P_k + P_k A_k & P_k B_k - C_k^T \\ B_k^T P_k - C_k & -D_k^T - D_k \end{bmatrix} \leq 0.
\]

3) Generalized plant for the optimization of tyre grip: In this section we will focus on a single aspect of performance, namely the tyre grip which is related to the tyre normal loads. We will use the r.m.s dynamic tyre load parameter \( J_3 \) [40] for a standard stochastic road profile given by
\[
J_3 = 2\pi \sqrt{\kappa V} |s^{-1} T_z \ast k(z_0 - z_r)(j\omega)|_2,
\]
where \( \kappa \) is a road roughness parameter and \( V \) the vehicle velocity.

We now calculate the generalized plant, \( G_{J_3}(s) \), corresponding to the block diagram of Fig. 9 and the performance measure \( J_3 \). The performance output corresponding to \( J_3 \) is given by \( z = \int k(z_0 - z_r) \) and the excitation input is the road disturbance signal \( w = z_r \). The measurement signal for the controller is the relative velocity of the suspension, \( \dot{z}_r - \dot{z} \), and the controller output is the suspension force \( F \). It was shown in [29] that,
\[
G_{J_3}(s) = \begin{bmatrix} 0 & -\frac{k}{m} & 0 & \frac{k}{m} \\ 1 & 0 & 0 & 0 \\ 0 & \frac{k}{m} & 0 & -\frac{k + k_t}{m} \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{m} \\ 1 & 0 \end{bmatrix}.
\]

Given a controller \( K(s) \) of order \( n_k \), with state-space representation as in (3), let the state-space representation of the closed-loop system resulting from the interconnection of the generalized plant \( G_{J_3}(s) \) and the controller be given by:
\[
\begin{bmatrix} \dot{x} \\ \dot{x}_k \\ \dot{z}_u - z_r \end{bmatrix} = \begin{bmatrix} A_{cl} & B_{cl} \\ C_{cl} & 0 \end{bmatrix} \begin{bmatrix} x \\ x_k \\ z_r \end{bmatrix}.
\]

Theorem 5: There exists a strictly positive-real controller \( K(s) \) of order \( n_k \) such that \( J_3 < 2\pi \sqrt{\kappa V} k_0 V \) and \( A_{cl} \) is stable, if and only if the following matrix inequality problem is feasible for some \( X_{cl} > 0, X_k > 0, Q, \nu^2 \) and \( A_k, B_k, C_k, D_k \) of compatible dimensions:
\[
\begin{bmatrix} A_{cl}^T X_{cl} + X_{cl} A_{cl} & X_{cl} B_{cl} \\ B_{cl}^T X_{cl} - I \end{bmatrix} < 0, \quad \begin{bmatrix} X_{cl} & C_{cl}^T \\ C_{cl} & Q \end{bmatrix} > 0,
\]
\[
\text{tr}(Q) < \nu^2, \quad \begin{bmatrix} A_k^T X_k + X_k A_k & X_k B_k - C_k^T \\ B_k^T X_k - C_k & -D_k^T - D_k \end{bmatrix} < 0.
\]
The first three LMIs are necessary and sufficient conditions for the existence of a stabilising controller that achieves an upper bound of \( r \) on the closed-loop \( \mathcal{H}_2 \)-norm [34]. The fourth LMI further restricts the controller to be strictly positive-real. Without the positive-real constraint the \( \mathcal{H}_2 \)-synthesis problem can be formulated as an LMI problem as shown in [34]. With the positive-real constraint it is not obvious how to do so, hence an iterative optimization method is employed to solve the Bilinear Matrix Inequality (BMI) problem locally. The method, which is described in [18], is to linearise the BMI using a first-order perturbation approximation, and then iteratively compute a perturbation that ‘slightly’ improves the controller performance by solving an LMI problem. The proposed scheme was implemented in YALMIP [22], which is a MATLAB toolbox for rapid prototyping of optimization problems. A feasible starting point must be given to the algorithm.

4) Tyre grip optimization results: The optimization of the \( J_3 \) measure was attempted in [40] over various fixed structure suspensions (see Fig. 10). In contrast, the iterative algorithm implemented in YALMIP was used to optimize \( J_3 \) over general second-order admittances \( K(s) \) in order to investigate whether \( J_3 \) can be improved further. The optimization was performed for \( k_c \) ranging from \( 1 \times 10^4 \) N/m to \( 12 \times 10^4 \) N/m in steps of 2000 N/m. The comparison of the optimization results obtained with YALMIP with those obtained by fixed-structure optimization are presented in Fig. 11.

The optimization results obtained with YALMIP are presented as three distinct curves suggesting that the structure of the suspension changes as the static stiffness varies. At low and high stiffness the YALMIP second-order admittance can do better than both the second-order S5 layout and the third-order S7 layout. An encouraging feature of the optimization algorithm is that it allows the change in the structure of the admittance as the static stiffness varies in order to obtain the minimum value of \( J_3 \). In the intermediate range \( K(s) \) turns out to be the network S10 shown in Fig. 12 consisting of an inerter, damper and spring in series [29].

**Figure 11.** Comparison of YALMIP optimization results with fixed-structure optimization results for \( J_3 \). (See Figure 10 for the configurations.)

**Figure 12.** Additional passive suspension networks incorporating springs, dampers, and inerters (a) S9 and (b) S10.

B. Analytical Solutions for Optimal Ride Comfort and Tyre Grip

The approaches of [29], [40] both require extensive numerical optimizations. The question whether the solutions obtained are global optima is not rigorously settled. Also, if a new set of vehicle parameters is chosen, the
Passive suspensions provide a simpler and cheaper means of suspension design and construction at the expense of performance limitations than active suspensions.

Numerical optimizations must be repeated. In [33] both of these issues are addressed for ride comfort and tyre grip performance measures in a quarter-car vehicle model. Six suspension networks of fixed structure are selected: S1–S4 in Fig. 10 and S9–S10 in Fig. 12. Global optima are derived as a function of the quarter-car model parameters. The optima are also parameterised in terms of suspension static stiffness, which can therefore be adjusted to approximately take account of other performance requirements, such as suspension deflection and handling.

1) The quarter-car model and suspension networks: We consider again the quarter-car model described in Fig. 8, where \( Y(s) \) is the admittance of one of the candidate suspension networks.

Network S1 models a conventional parallel spring-damper suspension and S2 contains a “relaxation spring” in series with the damper. S3, S4, S9 and S10 show extensions incorporating an inerter and possibly one “centring spring” (cf. [40]) across the damper. The mechanical admittance \( Y(s) \) for three of these layouts (S3, S9, S10) is now given for illustration:

\[
Y_5 = \frac{K}{s} + c + sb,
\]

where the matrix \( L \) is the unique solution of the Lyapunov equation

\[
AL + LA^T + BB^T = 0. \tag{5}
\]

The matrix \( L \) is then determined from the linear equations in (5) and the performance measures are then given by

\[
J_1 = 2\pi \sqrt{K\kappa |s^{-1}T_{2,\mu}(j\omega)|_2},
\]

where the results for a mixed performance of \( J_1 \) and \( J_3 \) are beneficial.

2) Performance measures and analytical expression: In addition to the r.m.s. dynamic tyre load parameter \( J_3 \) defined in (4) we also consider a ride comfort measure. This is the r.m.s. body acceleration in response to a standard stochastic road profile and is equal to

\[
J_1 = 2\pi \sqrt{K\kappa |s^{-1}T_{2,\mu}(j\omega)|_2}.
\]

See [40] for detailed derivations of the performance measures.

An analytical expression of the \( \mathcal{H}_2 \)-norm of the (stable) transfer function \( G(s) \) can be computed from a minimal state-space realization as

\[
|G|_2 = |C(sI - A)^{-1}B|_2 = (CLC^*)^{1/2},
\]

where the matrix \( L \) is the unique solution of the Lyapunov equation

\[
AL + LA^T + BB^T = 0. \tag{5}
\]

The matrix \( L \) is then determined from the linear equations in (5) and the performance measures are then given by

\[
J_1 = 2\pi \sqrt{K\kappa H_{iSp}},
\]

where \( H = CLC^T \) and \( i \) indicates the performance measure index and \( j \) the suspension network number.

3) Optimal solutions for mixed performance of \( J_1 \) and \( J_3 \): Optimal performance solutions for \( J_1 \) and \( J_3 \) individually and for suspension networks S1–S4, S9 and S10 have been computed in [33]. Furthermore, it is also important to consider combined optimal vehicle performance across different measures. Here we present the results for a mixed \( J_1 \) and \( J_3 \) measure:

\[
Y_9 = \frac{K}{s} + \left( \frac{s}{k+sc} + \frac{1}{sb} \right)^{-1},
\]

\[
Y_{10} = \frac{K}{s} + \left( \frac{s}{k} + \frac{1}{c} + \frac{1}{sb} \right)^{-1}.
\]
Efficiency of realization, as defined by the number of elements used, is much more important for mechanical networks than electrical networks.

\[ H_{1:2:S} = (1 - \alpha)m_s^2H_{1:S} + \alpha H_{3:S}, \]  

where \( \alpha \in [0,1] \) is a weighting between \( J_1 \) and \( J_3 \). The scaling factor \( m_s^2 \) is inserted to approximately normalise the measures and simplify the resulting formulae. Eq. (6) can be optimized with respect to the suspension parameters \([33]\). The resulting optimal solutions are drawn for a particular \( m_u, m_s \) and \( k_i \) in Fig. 13. In general it can be seen that networks involving inerters (especially S9 and S10) offer performance advantages over conventional networks for both \( J_1 \) and \( J_3 \) combined. The results also show that ride comfort \( (J_1) \) deteriorates as suspension static stiffness increases, and that tyre grip improves as unsprung mass is decreased, for all suspension networks.

5. Simulation-Based, Optimal Design of Passive Vehicle Suspensions Involving Inerters

In this section we will present the development of a simulation-based methodology for the analysis and optimal design of nonlinear passive vehicle suspensions. The methodology makes use of a nonlinear vehicle model which is constructed in the Matlab/Simulink toolbox SimMechanics. The vehicle model is in a 4-post rig configuration and it allows the detailed representation of the suspension geometry and the nonlinearities of the suspension elements. Several aspects of suspension performance are considered such as ride comfort, tyre grip and handling. For each aspect of performance we will propose time-domain performance measures that are evaluated after a simulation run. For the ride comfort and tyre grip performance we define appropriate road disturbance inputs and for the handling performance we define appropriate torque disturbances acting on the sprung mass. The results demonstrate the performance improvements which can be achieved using inerters over a conventional arrangement using nonlinear dampers.

A. Nonlinear Vehicle Model

The nonlinear vehicle model considered in this study is typical of a high-performance sports car with a fairly accurate description of the suspension geometry and the characteristics of the suspension elements. The approximate parameters used for the vehicle model are given by its sprung mass \( m_s = 1500 \) kg and its moments of inertia about its roll, pitch and yaw axes respectively \( (I_x = 400 \text{ kgm}^2, I_y = 2300 \text{ kgm}^2, I_z = 2500 \text{ kgm}^2) \), the front unsprung masses each with a mass of \( m_u = 50 \) kg, and the rear unsprung masses each with a mass of \( m_u = 55 \) kg. Both the front and rear suspensions are of a double wishbone arrangement with a front static stiffness of 55 kN/m and a rear static stiffness of 50 kN/m. The tyres are modelled as vertical springs of stiffness 350 kN/m (rear) and 320 kN/m (front). Both the front and rear suspensions are a parallel arrangement of a spring with a nonlinear damper. The nonlinear dampers have a dual rate characteristic with a smooth transition between the hard and soft settings. Such a dual-rate damper characteristic has been found to provide better combined performance in ride comfort and handling than a linear damper \([30]\). A static view of the animation of the vehicle model is shown in Fig. 14 in its nominal state, i.e. with no external disturbances applied to it.

B. Definition of Disturbances

For the evaluation of the ride comfort and tyre grip we use a kerbstrike road profile. The kerbstrike has height \( h_0 \), length 1 m, and transition ramps of unity slope. Let \( v \) be the speed of travel of the vehicle and \( y \) the height of the kerb. Then we have:

\[
y(t) = vt, \quad 0 < t \leq \frac{h_0}{v}, \]
\[
y(t) = \frac{h_0}{v} - \frac{1 - h_0}{v}t, \quad \frac{h_0}{v} < t \leq \frac{1}{u}.
\]
As well as offering new possibilities for “passive mechanical control” in a variety of applications, the inerter brought out strong connections with the classical theory of electrical circuit synthesis, reviving old questions and suggesting new ones.

\[ y(t) = 1 - vt, \quad \frac{1 - h_0}{v} < t \leq \frac{1}{v}. \]

The kerbstrike initially appears at the front left wheel and subsequently at the rear left wheel delayed by \( L/v \) seconds, where \( L \) is the wheelbase of the vehicle.

The load disturbances used for the assessment of handling are pitch and roll step signals applied on the sprung mass. Due to the left-right symmetry of the vehicle model the roll disturbance is a step about the roll axis that results in a negative roll angle of the sprung mass for some fixed time and then the step is removed so that the sprung mass recovers zero roll angle. The disturbance about the pitch axis is chosen such that it results in both pitching-up and pitching-down of the sprung mass since there is no front-rear symmetry. Again with the removal of the pitch disturbance the vehicle pitch angle is restored to zero. The actual magnitude of the pitch and roll disturbances is specified accordingly by taking into account the relative importance of the handling performance over the ride comfort and tyre grip performance.

**C. Definition of Performance Measures**

The performance measure for the ride comfort considers the weighted accelerations of the sprung mass, namely the heave \((\ddot{z})\), pitch \((\dot{\varphi})\) and roll \((\dot{\vartheta})\) accelerations. The acceleration weights are taken from [5] and represent discomfort felt by humans due to mechanical vibrations. The performance measure for tyre grip considers the tyre forces at the four wheel stations. The time-domain measures for ride comfort and tyre grip are defined as:

\[ J_{Rt} = \sqrt{\text{trace}(\dot{z}^T \ddot{z})} y_{-\text{kerbstrike}}, \]
\[ J_{Rt} = \sqrt{\text{trace}(F^T F)} y_{-\text{kerbstrike}}, \]

where the signal \( z_{aw} = [\ddot{z}, \dot{\varphi}, \dot{\vartheta}] \) denotes the weighted acceleration responses of the sprung mass, \( y \) denotes the road elevations at the four wheel stations and \( F \) denotes the tyre forces. It is easy to see that

\[ J_{Rt} = \sqrt{\dot{z}^T \ddot{z} + \dot{\varphi}^T \dot{\varphi} + \dot{\vartheta}^T \dot{\vartheta}} \]

so it represents the square root of the sum of the energies squared of the relevant signals. In the case of the kerbstrike disturbance the resulting signals are finite energy signals.

In order to define the time-domain handling measures we assume that we know the desired handling responses of the vehicle in the pitch and roll channels, both in bump and rebound in case they are different. The energy of the error (possibly weighted) between the actual and the desired response can then be used as a time-domain handling measure.

If the energy of the error is small then the handling of the vehicle is close to the desired handling performance. The time-domain handling measure is defined as:

\[ H_t = \sqrt{e^T \ddot{z}^T e + e^T \dot{\varphi}^T \dot{\varphi}}, \quad (7) \]

where \( e_{\text{roll}} \) is the error signal due to the application of the roll disturbance and \( e_{\text{pitch}} \) is the error signal due to the application of the pitch disturbance.

**D. Optimal Design of Nonlinear Suspensions**

In this section we use the nonlinear simulation model and the

![Figure 15. The new suspension network and the admittance function of the linear series connection of the spring, damper and inerter.](image-url)
defined performance measures to design a suspension network involving nonlinear dampers and inerters in order to improve the ride comfort, tyre grip and handling when compared to the performance achieved with the default nonlinear damper characteristic. The approach was to use the default network topology of the parallel combination of the spring and the nonlinear damper with an extra parallel network consisting of a series connection of a spring, a damper and an inerter as shown in Fig. 15. The cost function

$$J = \frac{1}{2} J_{bc} + \frac{1}{2} J_{bc}$$

was optimized over the front/rear soft settings of the nonlinear dampers and the front/rear parameters of the series network, where the subscript “0” denotes the performance of the default suspension. The following values were obtained after optimization:

$$J_{bc} = 0.98, \quad J_{bc} = 0.945, \quad H_{bc} = 1.003.$$ 

The above results indicate that the tyre grip is improved by 5% without deteriorating the ride comfort and handling performances. It is expected that including the hard settings of the nonlinear dampers as decision variables in the optimization and also using a cost function that includes all aspects of performance will also result in an improvement of the handling performance.

6. Play in mechanical networks with inerters

A physical inerter as shown in Fig. 3 contains mechanical play or backlash in e.g. the rack and pinion mechanism which may affect the performance of the device, its closed-loop stability and its mechanical durability. This section addresses the mathematical modelling of passive mechanical networks including play and their physical accuracy. The results are based on [32] and have shown that the treatment of play as an input-output operator in mechanical networks leads to unsatisfactory solutions from a physical point of view. In contrast, a behavioural definition of play (ideal play) does not suffer from these objections and appears more reasonable from a physical point of view.

A. The Play Operator

A number of different play definitions have been proposed in the literature: the dead-zone (Fig. 16(a)) and hysteresis model (Fig. 16(b)) with the latter commonly used as a basis for a formal mathematical approach to play. Both definitions aim to describe an apparently well-defined phenomena and give rise to two different mathematical descriptions. This raises the question of which model, or indeed either, is more satisfactory?

The behaviour of the play operator in Fig. 16(b) can be expressed as a condition of three hybrid states. Here, the position of the piston ($z_1$) and cylinder ($z_2$) are considered to be the input and output (follower), respectively.

1. Engagement—extension: $z_2 = z_1 + \epsilon, \quad z_1 = z_2 \leq 0$,
2. Engagement—compression: $z_2 = z_1 - \epsilon, \quad z_1 = z_2 \geq 0$,
3. Disengagement: $|z_1 - z_2| < \epsilon, \quad \dot{z}_2 = 0$.

For a simple mechanical network incorporating the play operator ($H$) in series with a damper (Fig. 17) several properties can be identified that are unsatisfactory from a physical point of view, [32]:

- During disengagement the force through the play element is not necessarily zero.

![Figure 16.](image-url)

(a) Graph of dead-zone play model. (b) Graph of hysteresis play model.

![Figure 17.](image-url)

Damped harmonic oscillator network. The letter $i$ indicates the input and $f$ the follower.
The inerter is defined to be a two-terminal mechanical device such that the applied force at the terminals is proportional to the relative acceleration between them.

2. The solutions of the network equations depend on the choice of inertial frame, namely, the addition of a constant velocity to all states may change switching times or eliminate them altogether.

3. During engagement the force through the play element is not restricted in sign.

4. The behaviour of the network is not invariant to a switch of terminals of the play operator.

B. Ideal Play
Following the shortcomings of the above play operator, a behavioural definition for ideal play was proposed in [32] which does not suffer from this criticism. Consider a physical representation of play as shown in Fig. 18(a) where \( z_1, z_2 \) are the terminal positions and \( F \) is the equal and opposite force applied at the terminals. The ideal play is defined to be completely characterised by the following three states:

- (engagement—extension): \( z_2 - z_1 = \varepsilon, \quad F \leq 0 \),
- (engagement—compression): \( z_2 - z_1 = -\varepsilon, \quad F \geq 0 \),
- (disengagement): \( |z_2 - z_1| < \varepsilon, \quad F = 0 \).

Note that the definition is invariant to terminal reversal and by definition always admits a force through the device of appropriate sign (see Fig. 18(b) for the modelling symbol). Also, we note that this definition allows the mechanical network to maintain invariance to the choice of inertial frame, since the three states only depend on the difference between \( z_1 \) and \( z_2 \).

However, since the ideal play does not admit an input-output graph, mathematical properties like well-posedness and the exclusion of limit points of switching are arrived at by analysing individual transition scenarios, [32]. By means of the network example shown in Fig. 19, one can show that at engagement of play impulsive forces \( P \) may occur and multiple solutions are obtained. Energy is dissipated when \( 2P_0 < P \leq P_0 \) and energy increases when \( P > 2P_0 \).

In order to regain well-posedness and capture the range of solutions indicated in Fig. 20, the network in Fig. 19 was systematically extended by the addition of compliance springs and dampers. This highlights a connection with the work of Nordin et al. [25] who proposed a model for backlash which is equivalent to the semi-ideal model in Fig. 21. This model was shown to be effective in modelling the practical behaviour of inerter with play [26].

7. Conclusions
This paper has described the background and application of a newly introduced mechanical circuit
element, the inerter, from its origin in modelling and circuit synthesis through to a high-profile application in Formula One racing. The role of the inerter to make the analogy between electrical and mechanical circuits exact has been emphasised. From a practical point of view, the inerter allows the most general passive mechanical impedances to be synthesised, which is not possible using the traditional analogy in which the mass element is used. From a theoretical point of view, the subject of transformerless synthesis of one-port networks is reopened with some interesting new twists. Several application areas for the inerter have been highlighted. The paper has given a detailed treatment of the application of the inerter to vehicle suspensions and discussed the deviation from ideal behaviour of practical devices.

Michael Z.Q. Chen was born in Shanghai. He graduated from Nanyang Technological University, Singapore, in 2003 with a B.Eng. degree in Electrical and Electronic Engineering, and from Cambridge University in 2007 with a Ph.D. degree in Control Engineering. He is currently a Lecturer in the Department of Engineering at the University of Leicester, England. He is a Fellow of the Cambridge Philosophical Society, a Life Fellow of the Cambridge Overseas Trust, and a member of the IEEE. Since 2008, he has been an Associate Editor of the *IEE Journal B–Intelligent Devices & Systems* and a reviewer of the *IEEE Transactions on Circuits & Systems, Automatica, the International Journal of Adaptive Control & Signal Processing*, and the *Journal of Sound & Vibration*, amongst others. His research interests include: passive network synthesis, vehicle suspensions, complex networks, and statistical mechanics.

Christos Papageorgiou was born in Limassol, Cyprus. He graduated from the University of Cambridge, UK, in 1999 with an M.Eng./B.A degree in Electrical and Information Sciences and in 2004 with a Ph.D. degree in Control Engineering. He held positions as a research associate in the Control Group at Cambridge University, as a researcher in the Electrical and Computer Engineering Department of the University of Cyprus and as a research assistant in the Automatic Control Group of Linköping University. His research interests include vehicle suspension control, flight control design and clearance, modelling and identification of mechanical devices and the application of convex optimization in controller design and analysis.

Frank Scheibe was born in Bremen, Germany. He received the M.Eng. degree in Electrical and Electronic Engineering from Imperial College London in 2003, and the Ph.D. degree in Control Engineering from Cambridge University in 2008. In 2005 he worked for McLaren Racing Ltd and in 2007/08 was a Vehicle Dynamics Engineer with McLaren Automotive Ltd, Woking, England. He is currently a Research and Development Engineer with the BMW Group, Munich, Germany. His research interests include nonlinear mechanical systems, vehicle suspensions, and hybrid engine control.

Fu-Cheng Wang was born in Taipei, Taiwan, in 1968. He received the B.S. and M.S. degrees in mechanical engineering from National Taiwan University, Taipei, Taiwan, in 1990 and 1992, respectively, and the Ph.D. degree in control engineering from Cambridge University, Cambridge, U.K., in 2002. From 2001 to 2003 he worked as a Research Associate in the Control Group of the Engineering Department, University of Cambridge, U.K. Since 2003 he has been with the Control Group of Mechanical Engineering Department at National Taiwan University, in which he is now an Associate Professor. His research interests include robust control, inerter research, suspension control, fuel cell control, medical engineering and fuzzy systems.

Malcolm C. Smith 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. He returned to Cambridge University as a Lecturer in the Department of Engineering in 1990, became a Reader in 1997, and Professor in 2002. 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. Prof. Smith was a corecipent 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. He is a Fellow of the IEEE.

References
[41] B. D. H. Tellegen, Theorie der Wisselstromen. Groningen, the Nether-