The impacting cantilever: modal non-convergence and the importance of stiffness matching

John Melcher\(^1\), Alan R. Champneys\(^1\)
and David J. Wagg\(^2\)

\(^1\)Department of Engineering Mathematics, and \(^2\)Department of Mechanical Engineering, University of Bristol, Queen's Building, Bristol BS8 1TR, UK

The problem of an Euler–Bernoulli cantilever beam whose free end impacts with a point constraint is revisited from the point of view of modal analysis. It is shown that there is non-uniqueness of consistent impact laws for a given modal truncation. Moreover, taking an \(N\)-mode compliant, bilinear formulation and passing to the rigid limit leads to a sequence of impact models that does not converge as \(N \to \infty\). The dynamics of such truncated models are studied numerically and found to give rise to quite different dynamics depending on the number of degrees of freedom taken. The simulations are compared with results from simple experiments that show a propensity for multiple-tap dynamics, in which higher-order modes lead to rapidly cycling intermittent contact. The conclusion reached is that, to derive an accurate model, one needs to avoid the impact limit altogether, and take sufficiently many modes in the formulation to match the actual stiffness of the constraining stop.

1. Introduction

Mechanical systems with repeated impacts arise in a wide variety of applications. Examples include atomic force microscopy (AFM), radio-frequency switches, Braille printers, impact forging, percussive moling, backlash in gears, freeplay in control surfaces and numerous vibro-impact problems involving buzz, squeak and rattle. See di Bernardo et al. [1] and Brogliato [2] for general theory and numerous examples. A particular
motivation for this work is an attempt to understand the complex dynamical properties of tapping-mode AFM as observed elsewhere [3–7]. Here, a flexible micro-scale cantilever is excited at its clamped end so that a narrow tip at the free end can repeatedly impact with a sample in order to capture surface properties. Fundamental questions arise as to what is the ‘correct’ rational-mechanics formulation of the impact of a continuously flexible structure in order to faithfully capture its post-impact dynamics.

A commonly used formalism is that of rigid-body mechanics in which impact is instantaneous and lossy, with the energy dissipation being captured by a Newtonian coefficient of restitution (CoR). The study of such single-degree-of-freedom (1DOF) impact oscillators goes back to the Eastern European literature beginning in the 1950s (see [8,9] and references therein), culminating in the work of Peterka in the 1970s [10]. The theory was rediscovered in the West a decade or so later through the work of Shaw & Holmes [11] and by Thompson, to whom this volume is dedicated, together with his student Ghaffari [12]. Much information is now understood about the influence of grazing and chattering on global dynamics (see [13] and [1, ch. 1] for reviews).

Experimental investigation of impact oscillators at the macro-scale has shown qualitative agreement with the theory [14–18] (the last being a different Thompson, albeit a former PhD student from the same UCL group!). However, precise quantitative agreement is hampered by the fact that Newtonian restitution is a gross approximation to the true process of energy dissipation. In reality, dissipation occurs through excitation of internal degrees of freedom of the objects undergoing impact. In the case of a clamped cantilever undergoing intermittent contact at its free end, this problem is particularly acute because of the propensity of each impact to excite higher-order modes of the structure.

Turning to tapping-mode AFM, an often adopted modality is that of dynamic force spectroscopy in which the cantilever’s deflection is used as a proxy for the interaction force between the tip and the sample. Then, using a model-based approach, the signal is inverted to reconstruct the tip–sample force as a function of the tip displacement (see [6,19] and references therein). Material properties of the sample, e.g. its stiffness, viscosity, etc., can then in theory be inferred from the data by fitting the reconstructed tip–sample forces to a model. Although such a model-based inversion technique can be used to capture single coefficients like that of Newtonian restitution, such an approach typically does not take into account the response of higher modes of the cantilever beam. Such higher modes can cause multiple-tap dynamics, where a single impact is followed by several others [7], such as a drummer’s stick during a drum roll. Such dynamic behaviour and the question of how to model it is the chief motivation for this paper.

To be concrete, we shall consider the problem of a cantilever that is subject to spatially distributed forcing and either viscous or material damping, whose free end is allowed to make point contact with a stiff constraint. The central question is how to faithfully model such a situation using finite-dimensional impact mechanics. In particular, in such problems one often appeals to two separate limiting processes to derive a tractable reduced-order mathematical model for such a configuration. The first process is that of taking the infinite stiffness limit, to avoid the numerical problems associated with different time scales of contact and free motion. Indeed, in the 1DOF case, Nordmark [20] has shown theoretically and via simulation that the dynamics of a bilinear oscillator converges to that of the corresponding impact oscillator in the limit that one stiffness becomes infinite. The second limiting process is modal truncation. For a linear beam, a finite collection of modes can capture the response exactly, provided the forcing is entirely within the subspace spanned by those modes. But impact introduces strong nonlinearity, of a distributional kind, and one cannot appeal to such standard modal convergence theorems.

The rest of this paper is organized as follows. Section 2 introduces the model configuration to be studied. It is shown how, given a compliant constraint, modal decomposition and Galerkin projection can be used to formulate a reduced-order model up to any order. Section 3 goes on to study the same model in the rigid impact limit, with a given coefficient \( r \). Two proposed methods from the literature for posing an impact law for a given-order modal truncation are presented. Section 4 then revisits such impact laws by taking a succession of truncated models from §3 and
passing to the large stiffness limit. It is shown that this process does not converge. In §5 it is shown via numerical simulation that the observed dynamics of these impacting N-mode truncations depend strongly on the choice of N. Section 6 then presents experimental results for an impacting cantilever and shows how to optimally capture the observations using reduced-order models. Finally, §7 draws conclusions and considers practical implications for AFM operation.

2. Beam with compliant constraint

In this section and the next two, we shall undertake a theoretical analysis of a cantilever beam with an isolated point impact. As is common in the literature, we shall make several simplifying assumptions, in particular that damping occurs only through viscous drag and that impact occurs precisely at the end of the cantilever. In the simulations in §5, we shall relax the former assumption and include material damping, in order to get a better match with the experimental results in §6. The theory is, in principle, unaffected by more complex damping models or more general impact positions, but becomes rather cumbersome to present owing to modal coupling and lack of explicit solutions.

(a) Continuum model

Consider a uniform, Euler–Bernoulli cantilever beam with proportional damping, spatially extended external forcing and end contact with a compliant stop (figure 1). Under the usual assumptions of beam theory, after an appropriate rescaling of space and time, its dynamics can be represented by the dimensionless equation

\[ w_{tt}(s, t) + 2\alpha w_t(s, t) + w_{ssss}(s, t) = f(s, t), \]

such that

\[ s = 0: \quad w = w_s = 0 \]

and

\[ s = 1: \quad w_s = 0, \quad w_{sss} = \begin{cases} 0, & w \leq d_0 \\ \frac{k}{4}(d_0 - w), & w > d_0. \end{cases} \]
Here, $w$ is the scaled transverse deflection of the beam, $s \in [0, 1]$ is an axial coordinate, with $s = 0$ representing the clamped end and $s = 1$ the free end. Subscripts $s$ or $t$ on $w$ denote partial differentiation with respect to that variable, $2\alpha$ is a coefficient of mass-proportional viscous damping and $f(s, t)$ represents an axially distributed, time-varying loading. Note that the beam’s bending stiffness and mass density have both been scaled to unity in this formulation, without loss of generality. The presence of a compliant, one-sided amplitude constraint at the free end of the cantilever is captured via the piecewise-linear shear boundary condition, where $d_0$ is the equilibrium distance between the free end of the cantilever and the one-sided constraint, which (for convenience of later calculations) is assumed to have a scaled stiffness $k/4$.

The equation of motion (2.1) represents a piecewise-smooth boundary-value problem (BVP) that can be considered as a nonlinear dynamical system on an appropriate space of functions defined on the interval $[0, 1]$. The only nonlinearity is the Heaviside discontinuity in the boundary conditions that occurs when $w(1, t) = d_0$. But if we consider motion that does not cross this constraint, then (2.1) represents a linear, self-adjoint equation. In what follows, we shall let the subscript $+$ denote the region $w(1, t) \geq d_0$ where the beam is constrained by the stop, and $-$ denote the region $w(1, t) < d_0$ where the cantilever is unconstrained.

Natural frequencies for the unforced, undamped problem can be obtained in the $+$ and $-$ regions separately by restricting motion to the relevant region, setting $\alpha = f = 0$ in (2.1) and looking for solutions of the form $w(s, t) = \psi_i^\pm(s)e^{i\omega t}$. This produces a countably infinite set of mutually orthogonal eigenfunctions,

$$\psi_i^\pm(s) = \cos(\beta_i^\pm s) - \cosh(\beta_i^\pm s) - \sigma_i^\pm [\sin(\beta_i^\pm s) - \sinh(\beta_i^\pm s)],$$

(2.2)

where

$$\sigma_i^\pm = \frac{\cos \beta_i^\pm + \cosh \beta_i^\pm}{\sin \beta_i^\pm + \sinh \beta_i^\pm},$$

with $\beta_i^\pm$ being the $i$th positive solution to

$$\cos \beta \cosh \beta + 1 + \frac{k^\pm}{4\beta^3} (\sin \beta \cosh \beta - \cos \beta \sinh \beta) = 0.$$

(2.3)

Here, $k^\pm$ is a bimodal stiffness, with $k^- = 0$ and $k^+ = k$, and $(\beta_i^\pm)^2$ represents the $i$th eigenfrequency of the linear problem.

In the expression (2.3) the terms inside and outside the parentheses correspond to the characteristic equations for the eigenfrequencies of a clamped–pinned beam and a simple cantilever beam, respectively. The factor $k/(4\beta^3)$ represents a weighting between the two characteristic equations. If $k \ll (\beta^-)^3$, then expanding the solution of (2.3) in the $+$ region for small values of $k$ about $\beta_i^-$, we find

$$(\beta_i^+)^4 = (\beta_i^-)^4 + k + O(k^2),$$

(2.4)

where $k$ is the modal stiffness of the constraint for all of the free cantilever modes and $O$ denotes the order as $k \to 0$. Equation (2.4) shows that the first-order effect of $k$ is to stiffen the modes proportionally without affecting the unconstrained eigenfunctions. On the other hand, if $k \gg \beta_i^-$, then $\beta_i^+$ tends to the $i$th root of the characteristic equation of a clamped–pinned beam. Nevertheless, for all finite $k$, there exists an $i$ that is sufficiently large so that $k \ll \beta_i^- \approx \beta_i^+$.

(b) Modal truncation

Given the solution (2.2) to the eigenvalue problem for (2.1) in the $+$ and $-$ regions, it is tempting to pursue the development of a reduced-order, lumped-parameter model using a truncated eigenfunction expansion in each region separately. For a purely linear problem, we know by the eigenfunction expansion theorem [21] that such a procedure will converge. Here, however, there is a subtlety because the overall problem is nonlinear, and so one should not expect a purely linear eigenfunction expansion to correctly capture the solution.
To resolve this apparent paradox, suppose that the exact solution within each region, written in the form
\[ w(s, t) = \sum_{i=1}^{\infty} \psi_i^\pm(s) q_i^\pm(t), \]
is decomposed into a truncated series approximation and a residual error:
\[ w^N(s, t) = \sum_{i=1}^{N} \psi_i^\pm(s) q_i^\pm(t) \quad \text{and} \quad e^N(s, t) = \sum_{i=N+1}^{\infty} \psi_i^\pm(s) q_i^\pm(t). \]
Indeed, this was the approach adopted by Moon and Shaw [22], where the first mode of the cantilever and the clamped–pinned beam were retained in a lumped-parameter, 1DOF, reduced-order model for the dynamics of a cantilever in the presence of a rigid and one-sided constraint. The orthogonality relations of the modes ensure only that \( w^N(s, t) \) and \( e^N(s, t) \) are orthogonal to one another. But this is not the case for \( w^N(s, t) \) and \( e^N(s, t) \). Consequently, at each switching event between + and − regions, a portion of the exact solution (the residual) is lost when changing basis from the unconstrained to the constrained eigenfunctions or vice versa. Thus, when crossing from the + region to the −, some of the information in \( w^N \) is projected onto \( w^N \) when we change basis, and vice versa when crossing from − to +. Hence, if simulating for a long time, it is not clear \textit{a priori} that this form of modal truncation will converge as \( N \to \infty \), as errors can accumulate upon repeated crossing of the switching condition. In what follows, we shall refer to the eigenbasis \( \{ \psi_i^+ \} \) as the set of \textit{constrained} cantilever modes and \( \{ \psi_i^- \} \) as the \textit{unconstrained} cantilever modes. To simplify notation, we shall let \( \phi_i = (-1)^{i+1} \psi_i^- \), where the sign of \( \phi_i \) is chosen conveniently so that \( \phi_i(1) = 2 \), for all \( i \).

A more tractable approach for modal truncation, for which the study of convergence is more straightforward, is to use a single eigenbasis. That is, a reduced-order model is obtained through a Galerkin decomposition using the unconstrained cantilever eigenfunctions \( \{ \phi_i(s) \} \) as the trial basis. To do this, it is natural to incorporate the interaction force between the beam and the constraint directly in the partial differential equation (2.1) as an additional forcing term rather than through the boundary conditions. This can be achieved as follows:
\[ w_{tt}(s, t) + 2\alpha w_t(s, t) + w_{ssss}(s, t) = \delta(s - 1) h(w(1,t)) + f(s, t) \]
such that
\[ s = 0 : \quad w = w_s = 0 \]
and
\[ s = 1 : \quad w_{ss} = w_{ssss} = 0, \]
where
\[ h(d) = \begin{cases} 0, & d \leq d_0 \\ \frac{k}{4}(d - d_0), & d > d_0 \end{cases} \]

Now, we make a global decomposition
\[ w(s, t) = \sum_{i=1}^{\infty} \phi_i(s) \eta_i(t), \]
where the \( \phi_i \) are unconstrained eigenfunctions that are scaled such that the orthogonality relation
\[ \int_0^1 \phi_i(s) \phi_j(s) \, ds = \delta_{ij} \]
is enforced, with \( \delta_{ij} \) being the Kronecker delta function. Substitution of (2.7) into (2.5) and taking the inner product with \( \phi_j \) yields
\[ \ddot{\eta}_j(t) + 2\alpha \dot{\eta}_j(t) + k_j \eta_j(t) + \phi_j(1) \sum_{i=1}^{\infty} \phi_i(1) \eta_i(t) = \frac{k d_0}{4} \phi_j(1) + f_j(t), \]

\[ s = \{ \quad \]
Figure 2. Comparison between the constrained eigenfrequencies of the full BVP (dashed lines) and the constrained eigenfrequencies of the reduced-order model (solid lines) for (a) 1DOF, (b) 3DOF and (c) 10DOF reduced-order models. (d) Scaled representation of the highest eigenfrequency in (a–c). (Online version in colour.)

where

\[ f_j(t) = \int_0^1 \phi_j(s)f(s,t)\,ds \quad \text{and} \quad k_j = (\beta_j^-)^4, \]

and a dot is used to denote differentiation in time. Letting

\[ \eta(t) = [\eta_1(t), \eta_2(t), \eta_3(t), \ldots, \eta_N(t)]^T \quad \text{and} \quad \phi(s) = [\phi_1(s), \phi_2(s), \phi_3(s), \ldots, \phi_N(s)]^T, \]

an \( N \)-degrees-of-freedom (NDOF) reduced-order model may be expressed as

\[ \ddot{\eta}(t) + 2\alpha \dot{\eta}(t) + [K^\pm] \eta(t) = F^\pm(t), \quad (2.9) \]

where \( F^-(t) = [f_1(t), f_2(t), f_3(t), \ldots, f_N(t)]^T \),

\[ F^+(t) = F^-(t) + \frac{k d_0}{4} \phi(1), \quad K^-_{ij} = \delta_{ij} k_j \quad \text{and} \quad K^+_{ij} = \delta_{ij} k_j + k. \]

The switching surface between the + and − regions is now given by

\[ \phi^T(1)\eta(t) = d_0. \]

(c) Modal convergence

Figure 2 shows how the eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_N \) of the stiffness matrix \( [K^+] \) approximate the modal stiffnesses \((\beta_1^+)^4, (\beta_2^+)^4, \ldots, (\beta_N^+)^4\) of the original continuum model. Notice how, for each value of \( N \), the highest-order mode diverges from the exact solution as \( k \to \infty \). Recall that each
\( \beta_i^+ \) of the exact solution tends to the ith mode of the clamped–pinned beam, when \( k \gg (\beta_i^+)^3 \). In the reduced-order model, however, as \( k \) becomes large, the highest mode corresponds to in-phase contributions of the unconstrained cantilever modes with eigenvector

\[
v_N = [N^{-1/2}, N^{-1/2}, \ldots, N^{-1/2}]^T
\]

and eigenvalue \( \lambda_N = kN \). Therefore, if \( k \) exceeds \( O(\beta_N^+)^3 \), the highest eigenfrequency of the truncated model and the exact eigenfrequency diverge.

This non-convergence between eigenvalues of the truncated model and the BVP in the limit \( k \to \infty \) is in fact an unavoidable consequence of performing Galerkin discretization using the unconstrained eigenmodes, which satisfy natural (force- and moment-free) boundary conditions at the free end. In the limit that the stiffness of the constraint becomes large compared with that of the highest-order included eigenfunction, the stiffness of the constraint at the free end makes the boundary condition much closer to a geometric (pinned) one. Therefore, the unconstrained cantilever modes do not provide a good approximation to the true dynamics. Nevertheless, for a given constraint stiffness \( k \), choosing \( N \) such that \( O(\beta_N^+)^3 \gg k \) ensures that the modal description of the reduced-order model will accurately capture that of the full continuum BVP. In particular, for fixed \( k \), we have modal convergence as \( N \to \infty \).

### 3. The impact limit and modal restitution laws

A common approach in modelling the dynamics of a flexible beam with a stiff, one-sided constraint has been to consider the so-called impact limit, where the constraint becomes rigid. This approach imposes the classical Newtonian CoR law at a localized point on the continuous beam given by

\[
w_i(s, t_{I+}) = -r w_i(s, t_{I-}) \quad \text{and} \quad w(s, t_I) = d_0
\]

and

\[
w_i(s, t_{I+}) = w_i(s, t_{I-}), \quad s \neq s_I,
\]

where \( 0 < r < 1 \) is the coefficient of restitution, \( t_{I\pm} \) is the time instantaneously after (+) or before (−) an impact occurs and \( s = s_I \) is the location of the impact. In what follows, we shall assume that \( s_I = 1 \), so that impact occurs at the free end of the beam. Incorporating (3.1a) into a modal model has the general form

\[
\ddot{y}(t_{I+}) = -[R]\ddot{y}(t_{I-}),
\]

where \([R]\) is a modal CoR matrix that allows the modal energy to be redistributed post-impact. An NDOF, reduced-order impact model is established once the modal CoR matrix \([R]\) is determined.

#### (a) Restitution law via collocation

Wagg & Bishop [23] (see also [24]) introduced a straightforward way to determine \([R]\), by enforcing (3.1a) to hold at a set of \( N \) collocation points \( s_j, j = 1, 2, \ldots, N - 1 \), along the beam. The modal CoR matrix \([R]_C \) calculated via such an approach is simply

\[
[R]_C = \begin{bmatrix}
\phi^T(s_1) \\
\phi^T(s_2) \\
\vdots \\
\phi^T(s_{N-1}) \\
\phi^T(s_N)
\end{bmatrix}^{-1}
\begin{bmatrix}
-1 & 0 & \cdots & 0 & 0 \\
0 & -1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & -1 & 0 \\
0 & 0 & \cdots & 0 & r
\end{bmatrix}
\begin{bmatrix}
\phi^T(s_1) \\
\phi^T(s_2) \\
\vdots \\
\phi^T(s_{N-1}) \\
\phi^T(s_N)
\end{bmatrix}.
\]

(3.3)

Note, as pointed out in [25], that \([R]_C \) is non-unique owing to the arbitrariness in the choice of the \( s_j \). Nevertheless, if the \( s_j \) are chosen to be equally spaced along the beam, then, in the limit as \( N \) becomes large, one might expect \([R]_C \) to converge, making this approach somewhat analogous to finite element analysis.
(b) Impulse-response restitution law

An alternative method for determining the modal CoR, proposed in [25], is to consider the impulse response of the reduced-order model and determine the magnitude of the impulse required to enforce (3.1a). To obtain such a matrix, consider a cantilever subjected to an isolated impulsive point force

\[ w_{\text{in}}(s,t) + w_{\text{ss}}(s,t) = p_0 \delta(s - s_1)\delta(t), \]  

(3.4)

where \( p_0 \) is the magnitude of the impulse. Pursuing a solution to (3.4) via eigenfunction expansion involves writing \( w(s,t) = \sum_{i=1}^{\infty} \psi_i(s)q_i^-(t) \) and truncating the series after \( N \) terms. Choosing \( s_1 = 1 \), projecting the solution onto the \( j \)th mode and solving for \( q_j^-(t) \) yields

\[ q_j^-(t) = p_0 \psi_i^-(1)(\beta_j^-)^{-2}\sin(\beta_j^-)^2t, \quad t > 0. \]

Substitution of the truncated modal expansion for the impulse response into (3.1a) solves for the unknown \( p_0 \) and establishes the relationship

\[ \dot{q}_j^-(t_+) = \dot{q}_j^-(t_-) - \frac{r + 1}{N} \sum_{i=1}^{N} q_i^-(t_+). \]

Following this approach, the impulse-response CoR \( [R] = [R]_{\text{IR}} \) for the 2DOF, reduced-order model is given by

\[ [R]_{\text{IR}} = \frac{1}{N} \begin{bmatrix} r + 1 - N & r + 1 & \ldots & r + 1 \\ r + 1 & r + 1 - N & \ldots & r + 1 \\ \vdots & \vdots & \ddots & \vdots \\ r + 1 & r + 1 & r + 1 & r + 1 - N \end{bmatrix}. \]  

(3.5)

4. Restitution law via infinite stiffness limit

Rather than impose the CoR law in (3.1a) \textit{a priori}, we shall now seek to derive a general form of CoR by considering the modal compliant model in the limit that the constraint stiffness \( k \to \infty \).

To this end, suppose we let \( k = \varepsilon^{-2}\kappa \), where \( 0 < \varepsilon \ll 1 \) is a scaling parameter and \( \kappa = O(1) \) as \( \varepsilon \to 0 \). In this case it is straightforward to see that the time taken for a trajectory of the dynamical system to pass through the + region is \( O(\varepsilon^{-1}) \). Thus, if we take the limit \( \varepsilon \to 0^+ \), this time tends to zero and we will have an equivalent impact oscillator with a modal impact law (3.2). The question arises how the impact law \([R]_N\) derived by this asymptotic approach applied to the compliant \( N \)-modal truncation will correspond to either \([R]_C\) or \([R]_{\text{IR}}\). The convergence properties of \([R]_N\) as \( N \to \infty \) are also of interest. In order to answer these questions, we shall consider separately the cases \( N = 1 \) and \( N = 2 \), before tackling the general case. Finally, to simplify the analysis, we consider the special case of \( d_0 = 0 \).

(a) Single-degree-of-freedom impact oscillator

Consider the 1DOF compliant model with initial conditions on the switching surface and impending motion in the + region:

\[ \ddot{\eta}_1(t) + 2\alpha \dot{\eta}_1(t) + (k + \kappa)\eta_1(t) = f_1(t), \quad \eta_1(t) \geq 0, \quad \eta_1(0) = 0 \quad \text{and} \quad \dot{\eta}_1(0) = \varepsilon^{-1}v^+ > 0. \]  

(4.1)

Letting \( k = \varepsilon^{-2}\kappa, t = \varepsilon \tau \) and \( \xi = \eta_1 \), we can rewrite (4.1) as

\[ \ddot{\xi}_{\varepsilon} + 2\alpha \dot{\xi}_{\varepsilon} + \kappa \dot{\xi}_{\varepsilon} = O(\varepsilon^2), \quad \xi > 0, \quad \dot{\xi}(0) = 0 \quad \text{and} \quad \ddot{\xi}(0) = v^+. \]  

(4.2)

Substituting \( \xi = \xi^0 + \varepsilon \xi^1 + \ldots \) into (4.2) and collecting terms by order of \( \varepsilon \) yields

\[ O(\varepsilon^0): \xi^0_{\varepsilon} + \kappa \xi^0_{\varepsilon} = 0 \quad \text{and} \quad O(\varepsilon^1): \xi^1_{\varepsilon} + \kappa \xi^1_{\varepsilon} = -2\alpha \xi^0_{\varepsilon}, \]
which has solution

$$\xi(\tau) = v^+ \kappa^{-1/2} (1 - \varepsilon \alpha \tau) \sin \frac{1}{2} \kappa \tau + O(\varepsilon^2), \quad 0 \leq \tau \leq \Delta^+,$$

(4.3)

where

$$\Delta^+ = \pi \kappa^{-1/2}$$

(4.4)

is the time of passage in the $+$ region. Differentiating (4.3) with respect to time and substituting $\tau = \Delta^+$, we find

$$v^- = -r_e v^+, \quad \text{with} \quad r_e = 1 - \alpha \pi \kappa^{-1/2} \varepsilon,$$

(4.5)

where $v^-$ is the velocity when the oscillator re-enters the $-$ region. We therefore call $r_e$ the effective CoR.

Figure 3a shows a comparison between the return velocity $v^-$ predicted by the impact law (4.5) and that obtained by simulating the 1DOF model (4.1) for different values of the constraint stiffness. Specifically, letting $\kappa = \varepsilon^{-2}$, $\eta_1(0) = 0$ and $\dot{\eta}_1(0) = \varepsilon^{-1}$, the effective restitution coefficient $r_e^2$ is compared with the computed $(v^-/v^+)^2$, as $\varepsilon$ is varied. Note, from the figure, the close agreement for small $\varepsilon$. Also, note that $r_e \to 1$ in the limit as $\varepsilon \to 0^+$. However, in experiments, Thompson et al. [18] observed low values of the CoR that were attributed to energy transfer to higher modes during contact. Hence, we will next consider the impact limit of the 2DOF model.
(b) Two-degrees-of-freedom model

Next, we consider the effect of an additional mode. For a 2DOF reduced-order model to the + region, the governing equation reduces to

$$\ddot{\eta}(t) + 2\alpha \dot{\eta}(t) + [K^+] \eta(t) = F^+ (t) \quad \text{and} \quad \phi^T(1) \eta(t) > 0. \quad (4.6)$$

Substituting $k = \kappa \varepsilon^{-2}/2$ into $[K^+]$ and scaling by $\varepsilon^2$, we find the eigenvectors are

$$[V] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} + O(\varepsilon^2). \quad (4.7)$$

Letting $\eta(t) = [V] q(t)$, $q(t) = [q_1(t), q_2(t)]^T$ and $t = \varepsilon \tau$, (4.6) can be expressed as

$$q_{r1}(\tau) + 2\alpha \varepsilon q_r(\tau) + \begin{bmatrix} 0 & 0 \\ 0 & \kappa \end{bmatrix} q(\tau) = O(\varepsilon^2), \quad q_2(\tau) > 0. \quad (4.8)$$

Notice that $q_1$ is orthogonal to the switching surface. The lowest-order mode imitates a clamped-pinned beam mode with zero amplitude at the free end. Letting $\xi = q_2$, we recover (4.2) for the 1DOF oscillator. It follows that the velocity of $q_2$ obeys the impact law given in (4.5). Additionally, integrating the equation for $q_1$ and expanding in powers of $\varepsilon$, we relate $q_{1r}(0) = v_1^+$ to the return velocity $q_{1r}(\Delta^+) = v_1^-$ by

$$v_1^- = v_1^+(1 - 2\alpha \pi \kappa^{-1/2} \varepsilon) + O(\varepsilon^2). \quad (4.9)$$

Thus, the modal CoR is given by

$$[R]_2 = [V] \begin{bmatrix} 1 - 2r_e & 0 \\ 0 & r_e \end{bmatrix} [V]^T. \quad (4.10)$$

Substituting $[V]$ from (4.7) into (4.10) we find

$$[R]_2 = \frac{1}{2} \begin{bmatrix} r_e - 1 & 1 - 3r_e \\ 1 - 3r_e & r_e - 1 \end{bmatrix}. \quad (4.11)$$

Notice that $[R]_2$ deviates in the off-diagonal terms from the expression (3.5) for the impulse-response restitution law $[R]_{1R}$. The source of the discrepancy is that the energy loss captured in $r_e$ is from the external viscous damping, whereas $r$ in (3.5) captures the ad hoc energy dissipation modelled by the Newtonian impact law. We note that both converge in the limit as $r$ and $r_e$ tend to unity.

The dynamics of the 1DOF and 2DOF impact models are fundamentally different post-impact. To see this, consider pre-impact, initial conditions

$$\eta(0) = 0 \quad \text{and} \quad \dot{\eta}(t) = v^+ = [\varepsilon^{-1} v_1^+, 0]^T.$$

The oscillator returns to the $-$ region with a velocity

$$v^- = \varepsilon^{-1}[v_1^-, v_2^-]^T, \quad \text{where} \quad v_1^- = -R_{11} v_1^+ \quad \text{and} \quad v_2^- = -R_{21} v_1^+,$$

and $R_{11}$ and $R_{21}$ are the elements from $[R]_2$. Letting $k = \varepsilon^{-2}$, the post-impact velocities predicted by the impact law are compared with numerical simulations of the reduced-order, compliant model as $\varepsilon$ is varied in figure 3b,c. In the case of the 2DOF model, the effective CoR of the first mode tends to zero as $\varepsilon \to 0^+$. In other words, the motion of the first mode of the 2DOF reduced-order model in the impact limit becomes completely uncorrelated with the motion of the equivalent 1DOF model after the first impact.

(c) General $N$-degrees-of-freedom case

For the NDOF case, let $\kappa = kN \varepsilon^2$. Following the approach for the 2DOF case, we substitute $q(t) = [V]^T \eta(t)$ to diagonalize (2.9) in the $+$ region. Without solving for either $[V]$ or $[A]$, we know that
\[ \lambda_j = \mathcal{O}(\varepsilon^2) \text{ for } j < N \text{ and } \lambda_N = \kappa + \mathcal{O}(\varepsilon^2) \text{ and the eigenvectors are mutually orthogonal with the eigenvector of the high-frequency mode given by} \]

\[ \mathbf{v}_N = [N^{-1/2}, N^{-1/2}, \ldots, N^{-1/2}]^T. \]  

(4.12)

It follows that the equation of motion in the + region can be expressed as

\[ \mathbf{q}_e(t) + 2\alpha\varepsilon\mathbf{q}_e(t) + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \kappa \end{bmatrix} \mathbf{q}(t) = \mathcal{O}(\varepsilon^2), \quad q_N(t) > 0. \]  

(4.13)

Substituting \( \xi = q_N \) into the equation for the \( N \)th mode in (4.13), we again recover (4.2). Accordingly, the time of passage in the + region and the return velocity of the \( N \)th mode are given by (4.4) and (4.5), respectively. The equations for \( q_j, j < N, \) are again over-damped oscillators for which the return velocities are predicted by (4.9). It follows that the NDOF modal CoR is given by

\[ [R]_N = [V] \begin{bmatrix} 1 - 2r_e & 0 & \cdots & 0 \\ 0 & 1 - 2r_e & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & r_e \end{bmatrix} [V]^T. \]  

(4.14)

From (4.12) and given that \( [V] \) is an orthogonal matrix, it is possible to show that

\[ [R]_N = \begin{bmatrix} 1 - 2r_e + \frac{3r_e - 1}{N} & \frac{3r_e - 1}{N} & \cdots & \frac{3r_e - 1}{N} \\ \frac{3r_e - 1}{N} & 1 - 2r_e + \frac{3r_e - 1}{N} & \cdots & \frac{3r_e - 1}{N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{3r_e - 1}{N} & \frac{3r_e - 1}{N} & \frac{3r_e - 1}{N} & 1 - 2r_e + \frac{3r_e - 1}{N} \end{bmatrix}. \]  

(4.15)

Substituting \( N = 1 \) and \( N = 2 \) recovers the results for the 1DOF and 2DOF oscillators, respectively.

\( \text{(d) Modal non-convergence as } N \to \infty \)

Modal non-convergence in the impact limit can be demonstrated readily with the NDOF modal CoR given in (4.15). The post-impact velocity profile is given by

\[ w^N_t(s, t_{1+}) = -\mathbf{\phi}^T(s)[R]_N\mathbf{v}^- . \]  

(4.16)

For the truncated model to converge, \( w^N_t(s, t_{1+}) \) must converge for a given \( \mathbf{v}^- \) as \( N \to \infty \). However, increasing \( N \) in the \( [R]_N \) excites higher-order modes to the same degree as the lower-order modes.

Figure 3d shows the truncated approximation for the post-impact velocity distribution \( w^N_t(s, 0^+) \) for \( N = 1, 2, 3, 10, r_e = 1 \) and \( \mathbf{v}^- = [1, 0, 0, \ldots, 0]^T \). Clearly, the post-impact velocity does not converge as modes are added to the restitution law. This non-convergence comes from taking an \( N \)-modal model, passing to the impact limit \( \varepsilon \to 0 \), and then considering what happens as \( N \to \infty \). It would appear that these two separate limiting processes are incommensurate.

To gain further insight, it is instructive to consider again the results in figure 2. This shows that modal convergence in the impact limit is achieved only if the limits \( \varepsilon \sim k^{-1/2} \to 0 \) and \( N \to \infty \) are taken together. That is, for each \( N \) there is a maximum value of \( k = \mathcal{O}(\beta^2_N)^3 \) for which the model gives answers that are commensurate with other modal truncations. However, modelling the interaction with the constraint with an impact law also requires \( k \gg \mathcal{O}(\beta^2_N)^4 \). At least for a
uniform cantilever beam, the two conditions are incompatible. Thus, to accurately model the interaction between the cantilever and the constraint in a truncated model requires a compliant formulation.

5. Numerical studies

In the previous section, we found that for the eigenfrequencies of the reduced-order model to converge to the exact solution of the continuous BVP required $k \ll O(\beta_N^+)^3$. However, because the system is nonlinear, this requirement is only a necessary condition for convergence. As $k$ becomes large the excursion of the oscillator into the $+$ region is characterized by increasingly small length scales, where contributions from higher-order modes may not be considered negligible. To study convergence of the nonlinear system, we turn to numerical simulations. Numerical simulations using the model studied so far led to qualitatively similar results to those presented in what follows (although there was a quantitative difference in the amount of damping seen in the higher modes). However, to get an acceptable match with experimental results presented in the next section, we found it necessary to include material damping in addition to the viscous damping from the surrounding fluid.

(a) Inclusion of material damping

The simplifying assumption in the previous sections of linear, mass-proportional damping enabled two sets of uncoupled modes corresponding to the $+$ and $-$ regions, which made the analysis tractable. If we now add material damping to (2.5) we have

$$w_{tt}(s, t) + 2\alpha w_t(s, t) + \gamma w_{ssss}(s, t) + w_{ssss}(s, t) = \delta(s-1)h(1, t) + f(s, t) \quad (5.1)$$

such that $s = 0$: $w = w_s = 0$

and $s = 1$: $w_{ss} = w_{sss} = 0,$

where $\gamma$ is the coefficient of stiffness-proportional material damping and $h(\cdot)$ is given by (2.6). Again, using a trial basis $\phi_i(s)$ corresponding to the normal modes of the simple cantilever, the NDOF Galerkin model becomes

$$\ddot{\eta}(t) + [C] \dot{\eta}(t) + [K^\pm] \eta(t) = F^\pm(t), \quad (5.2)$$

where $C_{ij} = (2\alpha + \gamma k)\delta_{ij}$ is the modal damping attenuation and $K^\pm$ and $F^\pm$ have the same meaning as in (2.9).

While in the $-$ region, the damping operator can be written as a linear combination of the mass and stiffness operators. However, in the $+$ region, we find full coupling, which renders the perturbation analysis in the previous section less tractable. Nevertheless, we should expect similar conclusions.

On the other hand, (2.9) and (5.2) are equally amenable to numerical analysis. Representing (5.2) in state-space form gives

$$\dot{x}(t) = [A^\pm] x + \hat{F}^\pm(t), \quad (5.3)$$

where

$$x(t) = \begin{bmatrix} \eta(t) \\ \dot{\eta}(t) \end{bmatrix}, \quad [A^\pm] = \begin{bmatrix} [0] & [I] \\ -[K^\pm] & -[C] \end{bmatrix} \quad \text{and} \quad \hat{F}^\pm(t) = \begin{bmatrix} 0 \\ F^\pm(t) \end{bmatrix}.$$ 

The solution of (5.3) is

$$x^\pm(t) = \exp([A^\pm](t - t_0))x_0^\pm + \int_{t_0}^t \exp([A^\pm](t - t'))\hat{F}^\pm(t') \, dt', \quad (5.4)$$

where $x_0^\pm = x_0(t_0)$ is an initial condition either on the switching surface or, in the case that $f(t)$ is periodic with period $\omega$, on the stroboscopic Poincaré section $\{\omega t = 0 \mod 2\pi\}$. In the latter case,
$t_0$ can be reset to 0 without loss of generality. Note that the convolution integral in (5.4) can be evaluated by parts.

(b) Simulation via root finding

The difficulty in using the explicit solution (5.4) for simulation arises when solving for the times of passage $\Delta^\pm$ in the $+$ and $-$ regions, which are the roots of a transcendental equation. These can be found using a numerical root-finding method. To ensure that near-grazing events with the switching surface are detected, a golden-section search method is used to identify local extrema in the trajectory of $d(t)$ on the bounded interval $[t_0, 2\pi/\omega]$. The trajectory is then segmented such that $d(t)$ is monotonic on each segment with at most one root. The bisection method is used to find the root of the first segment to cross the switching surface.

For the numerical simulations, we choose the parameters $d_0 = 0$, $\alpha = 0$ and define $\zeta_1 = \gamma(\beta_1^{-2})/2 = 0.01$, which is the damping ratio of the fundamental mode. Additionally, we choose the specific form of harmonic forcing given by

$$f(x, t) = \frac{k_1 \psi_1(s)}{2\zeta_1 \psi_1(1)} \cos \omega t, \quad \text{with} \quad \omega = 2(\beta_1^{-2}), \quad (5.5)$$

which projects purely to the first mode, and has frequency set to the effective resonance of the fundamental mode in the impact limit [26]. The amplitude of $f(x, t)$ is chosen, without loss of generality, so that the unconstrained amplitude of $\omega(1, t)$ is unity, when $\omega = (\beta_1^{-2})$.

To simulate periodic orbits, we choose initial conditions $x^+(0) = 0$ and discard the first 1000 excitation cycles. We then say that the resulting trajectory is a period-$m$ orbit if $\|x(t_0 + 2\pi m/\omega) - x(t_0)\| < 1 \times 10^{-9}$. Such a periodic orbit with $n$ impacts is denoted $(m, n)$.

(c) Results

Figure 4 compares numerically simulated periodic orbits for the truncated, bilinear model with $N = 1, 2, 3, 10$ and $k = 100k_1 = \mathcal{O}(k_2)$. In each case, the above numerical procedure resulted in a unique periodic orbit. For the 1DOF model, this is a simple $(1, 1)$ orbit. However, for the 2DOF, 3DOF and 10DOF models there are additionally complexities due to the higher-order modes contributing significantly to the interaction between the cantilever and the constraint. Thus, the 2DOF model predicts a $(2, 4)$ orbit, whereas the 3DOF model has converged to a $(1, 3)$ orbit. Essentially, the first impact excites high-frequency modes that result in additional taps on the constraining surface. By contrast the 10DOF model converges to a $(1, 1)$ orbit with a complex interaction with the constraint. While $m$ and $n$ are equal for the 10DOF and 1DOF cases, quantitatively the 10DOF model is closer to the 3DOF model. The additional modes in the 10DOF model provide a mechanism for rapid energy dissipation during the interaction because the modal damping attenuation goes as $\beta_1^{-2}$ for the material damping model.

We now compare what happens to the corresponding bifurcation diagrams as we vary the constraint stiffness $k$, with the other parameters held at $\zeta_1 = 0.01$, $d_0 = 0$, $\alpha = 0$ and $\omega = 2(\beta_1^{-2})$. Figure 5 shows results of the computation of a ‘brute-force’ bifurcation diagram depicted in the impact $(d = d_0, \nu > 0)$ Poincaré section.

Note that for small $k$ ($k \leq \mathcal{O}(k_1)$) each model predicts a stable $(1, 1)$ periodic orbit with quantitative agreement. As $k$ approaches $\mathcal{O}(k_2)$ ($k_2 \approx 39.3k_1$), higher-order modes couple into the response during interaction with the constraint. These modes couple into the response much more rapidly than study of the linear convergence in §2 would suggest. That is, although the lowest-order modes have converged reasonably well for $k \gtrsim \mathcal{O}(k_2)$, the relatively high stiffness of the constraint allows small contributions from higher-order modes to significantly affect the overall interaction between the cantilever and the constraint.

Despite these considerations we can see from the results that $N = 3$ is sufficient to capture the qualitative features of the bifurcation, although, as we shall see, this is not universally true throughout the range of parameter values.
Figure 4. Simulated periodic orbits of the NDOF, bilinear oscillator (5.2) for $N = 1, 2, 3, 10$ in (a–d) respectively. Segments of the orbit correspond to the $-$ region and to the $+$ region (blue and red, respectively, online). Parameter values chosen are $k = 100k_1$, $\alpha = 0$ and $\gamma = 0.02(\beta_1^-)^2$, and the forcing function is given by (5.5). (Online version in colour.)

(d) Effective energetic coefficient of restitution via simulation

When conducting numerical simulations, one possible approach for deriving an equivalent impact model of low degrees of freedom is to compute an effective CoR by estimating the energy lost to higher-order modes through impact. We shall consider just such an approach to define a so-called energetic CoR $r_E$, which captures such energy lost from the fundamental mode in a 1DOF impact model (see [27,28] and references therein for related approaches in the literature).

Equating the energy lost during impact of a 1DOF impact oscillator to the energy lost from the fundamental mode of the NDOF bilinear oscillator in a single impact leads to an implicit definition of $r_E$ for that impact:

$$\frac{1}{2}(1-r_E^2)(v_1^+)^2 + \int_{t_i}^{t_i+\Delta} F_c(d(t'))\dot{\eta}_1(t') \, dt' = 0.$$  

(5.6)

Here, $t = t_i$ is the time instant at which passage occurs from the $-$ region to the $+$ region in the NDOF system, at which time the first mode has velocity $v_1^+$, $\Delta$ is the time of passage in the $+$ region, $\eta_j$ is $j$th modal coordinate of the NDOF bilinear oscillator and

$$F_c(d) = \begin{cases} 
0, & d \leq d_0 \\
\frac{k}{4}(d_0 - d)\phi_1(1), & d > d_0,
\end{cases}$$

is the interaction force between the beam and the constraint.
Figure 5. Monte Carlo bifurcation diagrams for NDOF truncated models with \( N = 1, 2, 3, 10 \) in (a–d), respectively, as \( k \) varies. All other parameters held as in figure 4.

We can then define a single energetic coefficient \( r_E \) for a specific set of parameters and initial conditions, assuming repeated impacts occur via:

\[
\frac{1}{2} (1 - r_E^2) \sum_{i=1}^{\infty} (v_i^{(i)})^2 + \lim_{t \to \infty} \int_{t_0}^{t} F_c[d(t')] \dot{\eta_1}(t') \, dt' = 0, \tag{5.7}
\]

where \( t_i \) and \( v_i^{(i)} \) are the set of impact times and corresponding velocities.

Note that there can be numerical difficulties in evaluation of (5.7). In the case that motion converges to a periodic orbit, the limit can be evaluated simply by taking \( t_0 \) to be a time after a transient behaviour has decayed, and evaluating the integral and sum over a single period. But, for an arbitrary set of parameters, there is no guarantee that the limit in (5.7) will converge, nor that the limit will be unique over different initial conditions. For instance, it is known that even 1DOF impact oscillators can undergo chattering sequences, in which there are an infinite number of impacts in a finite time, and can feature coexisting attractors for the same parameter values (see [1] and references therein).

Figure 6 shows a comparison between the dynamics of a 20DOF bilinear oscillator and the 1DOF impact oscillator with energetic CoR computed using the above approach. In both cases, the motion converges to a periodic orbit, and there is therefore a unique, computable, solution to (5.7).

In the case depicted in figure 6a, \( d_0 = 0.8 \), there is relatively large damping and parameters are chosen so that all initial conditions converge to a \((1, 1)\) orbit that is close to grazing. In this case, the computation (5.7) predicts \( r_E = 0.036 \) and there is a good quantitative agreement between the 20DOF model and the effective 1DOF impact oscillator. In this case, only a small portion of the total energy stored in the oscillator is redistributed to higher modes during impact, which is subsequently dissipated as the damping is sufficiently high. Nevertheless, \( r_E = 0.036 \) is significantly lower than that which would be predicted from a naive 1DOF model using a typical value of 0.5–0.8 for the CoR for macroscopic collisions of steel on steel.
Figure 6. Comparison between the 20DOF bilinear model for $k = 1000k_1$ and the 1DOF impact model. The solid and dashed lines correspond to the motion of the free end and the component from the first mode of the 20DOF bilinear model. The 1DOF impact model output is shown as a separate solid line (green online) in the $-$ region only. (a) $(1,1)$ orbits near grazing with $d_0 = 0.8$, $\omega = 1.15(\beta - 1)^2$ and $\zeta_1 = 0.1$. $k = 1000k_1$ for the bilinear oscillator and $r_E = 0.036$ for the impact oscillator. (b) Periodic orbits for $d_0 = 0$, $\omega = 2(\beta - 1)^2$ and $\zeta_1 = 0.01$ and $r_E = 0.74$. (Online version in colour.)

The case depicted in figure 6b is more like that in figure 4, in which $d_0 = 0$, and the damping is relatively weak. Here, (5.7) predicts $r_E = 0.74$, which is closer to what one would expect from macroscopic impact theory. However, the results show that there is neither good quantitative nor qualitative agreement between the 1DOF impact model and the 20DOF bilinear model. Two particular effects that are due to the higher-order modes are the ‘multiple-tap’ dynamics of the impact, where there are two separate intervals in the $-$ region, and the large additional oscillations in the dynamics, which are not damped out during the free motion away from the stop.

6. Experiments

In this section, we present the results obtained from the experimental rig depicted in figure 7. A mild steel cantilever beam with dimensions $300\,\text{mm} \times 25.5\,\text{mm} \times 1\,\text{mm}$ was mounted on a shaker table. A pin constraint was also mounted on a second cantilever beam with dimensions $300\,\text{mm} \times 25.5\,\text{mm} \times 12\,\text{mm}$. From the dimensions of the cantilever and the constraint, we estimate $k/k_1 = O(10^3)$. The natural frequency $\omega_1 = 2\pi \times 5.92\,\text{rad}\,\text{s}^{-1}$, and the damping ratio $\zeta_1 = 7.7 \times 10^{-4}$ of the fundamental mode was determined from a ring-down test. Velocity measurements at the point on the beam corresponding to the location of the constraint were made with a Polytec OFV-5000 laser Doppler vibrometer (LDV) in the ground frame. Owing to the low damping, the motion of the base end of the cantilever was found to be much smaller than that of its free end, so that flexural base motion could be neglected. The velocity signal from the LDV is sampled at 50 kHz and has a resolution of $4\,\mu\text{m}\,\text{s}^{-1}$.

In addition to the velocity measurement, a simple electronic circuit was used to accurately determine when the stop and the beam are in contact. The circuit consists of two resistors in series that are supplied with a constant voltage $V_0$. The cantilever beam and constraint act as a switch that shorts the first resistor when the beam is in contact with the constraint. The voltage over the second resistor was sampled at 100 kHz.

To excite the cantilever the shaker table is used to prescribe a harmonic motion to its base. The motion of the base imparts a uniformly distributed inertial force on the cantilever. The inertial force projects onto each eigenmode; however, the effect on higher eigenmodes is minimal because (i) the excitation frequency is closest to the fundamental mode, and (ii) the higher modes are substantially stiffer than the fundamental. Consequently, the forcing from the base excitation of the physical system can be approximated by (5.5).
Figure 7. (a) Impacting beam experimental set-up consisting of a flexible cantilever beam and constraint mounted on a shaker table. (b) A schematic of the impacting beam and measurement system. The velocity at the point of impact, which is near to the free end, is measured by an LDV mounted in the ground frame. A simple circuit is used to determine precisely when the beam is in contact with the constraint. (Online version in colour.)

Figure 8a shows the results of a particular experimental run over one period \( T = \frac{2\pi}{\omega} \) of the forcing, obtained after transients had been allowed to decay. The motion was found to be approximately periodic, and the presented results are typical of what happens during each period. Approximately, when viewed at a macro-scale, this is a (1, 1) periodic orbit. However, the detailed time series (shown in two successive zooms in figure 8b,c) reveal the motion actually to be more highly structured and to contain five separate time intervals in which the beam is in contact with the stop. There is also evidence of high-frequency, undamped motion between each of the impacts. It would be tempting to wonder if the high-frequency motion is due to measurement noise. However, a comparison between the excited motion of the beam without the stop in figure 8d reveals that the high-frequency response is several orders of magnitude above the noise threshold.

Figure 9 shows simulation results for the 20DOF bilinear model at the same parameter values as those estimated from the experimental data. Note the strong qualitative similarity with the experimental data: there is similar multiple-tap dynamics with five separate impacts, and significant energy in the higher modes.

7. Discussion

The key result of this paper has been to show that the two limiting processes of taking an infinite stiffness limit and an infinite number of modes are incommensurate. That is, one gets a different answer for reduced-order rigid models depending on whether mode truncation or the rigid impact limit is performed first. In fact, as we argued in §2, it was already known that there is non-uniqueness of impact laws consistent with Newtonian restitution if one takes modal truncations of a rigid model. (The situation is even worse if one includes friction in the impact, see e.g. [29,30] and references therein.) We have shown though that the process of taking a succession of modal truncations and passing to the rigid limit separately does not converge. Rather like a Gibbs phenomenon, the residual is never small. This observation is important because we have found in the simulations that the dynamics observed depends strongly on the number of modes taken in the truncation, with multi-modal models having the propensity to undergo drum-roll-like multiple-tap dynamics with a high-velocity impact followed by several low-velocity impacts in quick succession. The experimental results have shown that this latter form of dynamics occurs in reality.
**Figure 8.** Experimental data obtained from the impacting cantilever. (a) Velocity of signal near to the free end and (b,c) two successive zooms. Note that the data in (a) is passed through a low-pass filter for ease of depiction. A portion (blue online) of the time series represents unconstrained motion, whereas the other portion (red online) corresponds to motion in contact; dots represent time instants at which contact first occurs. (d) The power spectrum $S_v$ corresponding to this motion (upper, blue line), compared with the equivalent results in which the stop is not present (lower, green line). The theoretical unconstrained eigenfrequencies are marked on the $x$-axis. (Online version in colour.)

**Figure 9.** Simulation of the 200DOF, truncated, bilinear model with $\zeta_1 = \gamma (\beta - 1)^2 / 2 = 7.7 \times 10^{-4}$, $\alpha = 0$, $\omega = 1.7(\beta_1 - 1)^2$ and $k = 10^3 k_1$. (a) Velocity waveform for single period (meaning of curves as in figure 8). (b) Amplitude spectrum $V_n$ for the first 25 harmonics of the excitation frequency revealing contributions for the first four modes. The unconstrained eigenfrequencies are marked on the $x$-axis. (Online version in colour.)
Moreover, we have shown that attempts to define effective (energetic) restitution laws calculated by estimating the energy transferred to higher-order modes are at best fraught with difficulty, leading to coefficients of restitution that are dependent on the dynamics observed, which can be very different from that derivable from a macroscopic consideration. Moreover, in many dynamical regimes, especially where much energy is lost in impact and with low damping, this approach is woefully inaccurate.

The overall conclusion from this paper then is that passing to a rigid limit for a point contact is a dangerous thing to do. It is better to ascribe a large but finite stiffness to the constraint and then to take sufficiently many modes to resolve time scales up to and beyond the time scale of an individual contact.

The present results also have important implications for dynamic AFM, which aims to extract material property information from a sample substrate through the measured response of a microcantilever probe. In this respect, if the probe were to behave as a 1DOF oscillator, the impact limit would represent a singular point at which all material properties of the sample collapse into a single parameter, the CoR. Thus, in this limit, quantitative measurements of material properties, such as stiffness/elasticity, viscosity, etc., are not possible. However, the results in this paper show that, for an impacting cantilever beam, this limit is never reached because additional modes couple into the dynamics as the stiffness of the constraint is increased. This introduces the possibility of quantitative material property measurements over large variations in sample properties by measuring the response of higher-order modes of the microcantilever.

The authors acknowledge useful conversations with Arvind Raman, Oliver Payton and Chris Budd. The work of J.M. has been sponsored by the UK Engineering and Physical Sciences Research Council.

References


