Abstract—We consider the cantilever beam model proposed in the seminal paper of Sader [1], where fluid damping effects on an oscillating cantilever are modeled. We show that in Sader’s model, the transfer function of base-to-tip displacement fails to display high-frequency roll-off. It turns out that this is also true for the widely-used viscous damped beam model. We further demonstrate that the addition of Voigt damping to Sader’s model creates the high frequency roll-off expected of a realistic model.

I. INTRODUCTION

The atomic force microscope (AFM) [2] provides means for atomic-scale imaging. The AFM is capable of obtaining images with resolution many orders of magnitude higher than optical microscopes, which are bound by the optical diffraction limit at atomic scales. The AFM operates by moving a specimen or sample under a micro-cantilever beam with a sharp tip. The cantilever beam “feels” the topography of the sample surface based on the inter-atomic forces between the cantilever tip and the sample; see Fig. 1. There are two fundamental modes of operation for the AFM, static mode and dynamic mode.

In static mode operation [3] the base of the cantilever is not forced, and the cantilever tip is consistently at a small distance from the sample. It is known in this mode the cantilever tip and the imaged sample are subject to large lateral as well as vertical forces, which leads to considerable wear of the cantilever tip and the sample. This drawback of static mode operation renders it unsuitable for interrogating soft samples which are encountered in most biological matter.

In dynamic mode operation [4] the cantilever base is oscillated at a frequency close to its first resonant frequency. The changes in the frequency and amplitude of the cantilever oscillations, that result due to its interaction with the sample, are then monitored to infer sample properties. This method generally allows for a higher imaging resolution compared to the static mode, and has the additional advantage of being less intrusive on the sample. In this paper our focus will be on dynamic mode operation.

The prevalent cantilever models used in the AFM literature are those obtained by approximating the cantilever by a lumped system. In such cases the basic strategy is to use a finite-dimensional model that captures well the frequency response of the cantilever around its first resonant frequency. The disadvantage of this approach is that it ignores high frequency dynamics of the cantilever, which can lead to undesired behavior and spillover effects. Therefore in this paper we choose to work directly with the partial differential equations (PDEs) that describe a cantilever beam.

To obtain an accurate PDE description of the cantilever it is important to model dissipative and damping effects. Distributed damping mechanisms can be divided into two categories: internal and external. In the literature on beam theory, there is some dispute as to what constitutes the most accurate model of a beam’s internal damping mechanism; see [5, Chap. 8] and references therein. Structural damping [6], also known as Voigt or Rayleigh damping, is a commonly used internal damping model. A widely used external damping model is given by viscous damping [6], [7]. Sader [1] describes the external damping effects of the surrounding fluid on an oscillating cantilever beam. In this paper we consider cantilever beams with both internal and external damping, where for internal damping we use the Voigt model and for external damping we use the fluid model proposed by Sader.

We emphasize here the importance of accurately modeling the cantilever dynamics, and in particular, its damping mechanisms. It is true that once the cantilever is used in dynamic mode various nonlinear phenomena arise, which admittedly represent challenging research topics in their own right. However, a precise model of the cantilever itself is crucial to achieving higher imaging resolutions. This fact constitutes our main motivation for this work.

The main contributions of this paper are as follows. We first consider the fluid damping framework of Sader [1], which is a widely used framework in the physics literature, and we obtain the transfer function of base-to-tip displacement. Using a frequency domain analysis of this transfer function, we show that the fluid damping model of Sader leads to a cantilever model that is physically unrealistic, in that it does not display high-frequency roll-off. However, we show that the addition of Voigt internal damping remedies this problem and leads to a frequency response similar to that observed in experiments.

The insufficiency of external viscous damping as the only source of damping for the cantilever has already been pointed out in the context of the cantilever control problem [8], and our findings emphasize these results. Furthermore, we argue that it is important to consider the transfer function $F(s)$ of base-to-tip displacement because: (a) in most AFMs the base
The cantilever deflects due to atomic forces between the sample and the cantilever tip. The amount of deflection can be measured by a laser incident on the cantilever (not shown).

The physical meaning of the variables is as follows [9]: \( \psi \) is the displacement, \( \psi_x \) is the rotation, \( \psi_{xx} \) is the bending moment, and \( \psi_{xxx} \) is the shear force. The boundary condition \( \psi(0,t) = w(t) \) forces the base of the beam to move like \( w(t) \). The equations (1) are derived from the condition for the equilibrium of forces on differential beam segments [6]. A more elegant method of deriving the governing equations for a lossless beam (\( \alpha = 0 \)) is through the use of energy methods and Hamilton’s variational principle [10].

We emphasize that (1) is different from most beam models considered in the AFM literature in that, in addition to force interactions with the nonlinear spring at the tip of the beam, it also allows for displacement inputs at the base. This is indeed a natural setting for the dynamic mode AFM, where the base is generally forced with a known input and the tip interacts with a sample whose force effects on the tip are typically unknown.

A. Representation as Feedback Interconnection

By denoting
\[
  u(t) = \phi(y(t)),
\]
we rewrite system (1) in the following form
\[
\begin{align*}
  &\mu \psi_{tt} + \alpha EI \psi_{txxxx} + EI \psi_{xxxx} = f, \\
  &\psi(0,t) = w(t), \\
  &\psi_x(0,t) = \psi_{xx}(l,t) = 0, \\
  &\alpha EI \psi_{txxx}(l,t) + EI \psi_{xxx}(l,t) = \phi(y(l,t)),
\end{align*}
\]
with output
\[
y(t) = \psi(l,t).
\]

The boundary input to the beam is now \( v = [w' u]' \), where \( w \) is a displacement input and \( u \) is a force input. The advantage of the new formulation is that we can now use tools from feedback control theory to analyze the properties of the interconnection of system (2) with the nonlinearity \( \phi \).

B. Open-Loop Transfer Function

Assuming zero initial conditions, we perform a Laplace transform in time on (2) to get\(^1\)
\[
\begin{align*}
  &\psi_{xxxx} = -q(s)\psi, \quad 0 < x < l, \\
  &\psi(0,s) = w(s), \\
  &\psi_x(0,s) = \psi_{xx}(l,s) = 0, \\
  &\psi_{xxx}(l,s) = \frac{1}{EI(\alpha s + 1)} u(s),
\end{align*}
\]
with output
\[
y(s) = \psi(l,s),
\]
where
\[
  q(s) := \frac{\mu s^2 - \tilde{f}(s)}{EI(\alpha s + 1)}.
\]

\(^1\)To avoid clutter we use the same notation for \( \psi \) in the Laplace domain and in the time domain. The distinction will be clear from the context.
Note that in obtaining (3)–(4) we have used the following property of the Laplace transform
\[ f(x, t) \rightarrow f(s) \psi(x, s). \]
We henceforth use \( H(s) \) to denote the 1-by-2 transfer function from the boundary input \( v = [w \ u]^T \) to the output \( y \) of the linear system (3), i.e.,
\[ y(s) = H(s) v(s) = \begin{bmatrix} F(s) & G(s) \end{bmatrix} \begin{bmatrix} w(s) \\ u(s) \end{bmatrix}. \]
Using state-space methods, we obtain (see Appendix for details)
\begin{align*}
F(s) &= \frac{4 \cos(l \sqrt{q(s)/4}) \cosh(l \sqrt{q(s)/4})}{2 + \cos(l \sqrt{q(s)/4}) + \cosh(l \sqrt{q(s)/4})}, \\
G(s) &= \frac{\sin(l \sqrt{q(s)/4}) - \sinh(l \sqrt{q(s)/4})}{2 + \cos(l \sqrt{q(s)/4}) + \cosh(l \sqrt{q(s)/4})} \\
&\quad \times \frac{\sqrt{2}}{EI \sqrt{q(s)^3} (\alpha s + 1)}.
\end{align*}
Finally, it can be concluded from the analysis of [11] that the poles of the transfer functions \( F \) and \( G \) belong to the open left-half of the complex plane. Furthermore, exponentially stability of system (3) was established in [12, Chap. 4].

C. The Closed-Loop System

Once the nonlinearity \( \phi \) is put in place, the closed-loop system can be described by the following block diagram

\[
\begin{array}{ccc}
  w & \rightarrow & F \\
  \bar{w} & \rightarrow & G \\
  u & \rightarrow & y \\
\end{array}
\]

where \( w \) and \( \bar{w} \) denote displacement input at the base \( (x = 0) \) and force input at the tip \( (x = l) \), respectively.

There is another way to represent the above closed-loop system. By defining \( W(s) := F(s)/G(s) \) the previous block diagram can be redrawn as

\[
\begin{array}{ccc}
  w & \rightarrow & \bar{w} \\
  \bar{w} & \rightarrow & G \\
  u & \rightarrow & y \\
\end{array}
\]

where \( W \) can be interpreted as the transfer function from oscillation at the base to forcing at the tip.

III. BEAM MODEL OF SADER: FLUID DAMPING

In this section we first briefly describe the main results of Sader [1], where it is assumed that the beam has zero internal damping [i.e., \( \alpha = 0 \) in (2) and (3)] and all damping of the beam is provided by the surrounding fluid. Then we employ frequency domain analysis to show that the beam model of [1] does not yield a transfer function \( F(s) \) that demonstrates the high-frequency roll-off expected of physical systems.

Reference [1] considers a beam described by (2) with zero internal damping \( \alpha = 0 \), zero input at the boundaries \( w = u = 0 \), and distributed forcing \( f \) that accounts for the “hydrodynamic loading” due to the motion of fluid around the beam. Therefore \( f \) contains all “dissipation” effects and the beam is otherwise undamped. The surrounding fluid is assumed to be viscous and incompressible. Using equations of motion for the fluid, Sader shows that
\[ \tilde{f}(j\omega) = \kappa \omega^2 \Gamma(j\omega)^*, \quad \kappa = \frac{\pi}{4} d^2 \rho, \]
where \( \Gamma \) is a complex-valued function of frequency, \( \rho \) is the density of the fluid, \( d \) is the width of the beam (which is assumed to be much smaller that the length of the beam), and \( \Gamma^* \) is the complex conjugate of \( \Gamma \). The dependence of \( \Gamma \) on \( \omega \) involves modified Bessel functions of the third kind. More specifically
\[ \Gamma(\omega) = 1 + \frac{4 j K_1(-j \sqrt{R})}{\sqrt{j R K_0(-j \sqrt{R})}}, \quad R = \frac{d^2 \rho}{4 \eta \omega}, \]
where \( \eta \) is the viscosity of the fluid and \( \rho \) is its density. In particular
\[ \Gamma(\omega) \rightarrow 1 \quad \text{as} \quad \omega \rightarrow \infty. \]

Fig. 2 shows the magnitude of the imaginary and real parts of \( \Gamma^* \) for the parameter values \( d = 30 \times 10^{-6} \text{m}, \rho_{\text{air}} = 1.16 \text{kg/m}^3, \) and \( \eta_{\text{air}} = 18.6 \times 10^{-6} \text{Pa.s} \). The imaginary part of \( \kappa \Gamma^* \) models the dissipative effects of the surrounding fluid, whereas the real part of \( \kappa \Gamma^* \) acts as added mass to the beam [1]. In other words, a beam with mass per unit length equal to \( \mu \), when immersed in fluid, behaves like a beam of mass per unit length equal to \( \mu + \kappa \text{Re}(\Gamma^*) \) in vacuum.

Let us now consider the fluid damping model of Sader in conjunction with nonzero boundary inputs at the base and tip of the beam, \( w \neq 0, u \neq 0 \). Henceforth we will refer to this system as “Sader’s model”. From (4) and (7) it is clear that the transfer functions \( F(j\omega) \) and \( G(j\omega) \) corresponding...
to Sader’s fluid-damped beam can be found from (5) and (6) via the substitution
\[ q(j\omega) = -\frac{\omega^2}{EI} (\mu + \kappa \Gamma(\omega))^\ast. \]

Fig. 3 shows the magnitude plots of \( F(j\omega) \) (red) and \( G(j\omega) \) (blue) for system (10) and (3) it is clear that the transfer functions \( F(j\omega) \) and \( G(j\omega) \) corresponding to system (10) can like an undamped system as the frequency of oscillations is increased. This is clearly displayed in Fig. 4.

IV. BEAM WITH BOTH FLUID AND VOIGT DAMPING

In this section we consider a beam model that has both fluid and internal damping. A model of the beam with fluid damping has been proposed by Sader [1] and was described in some detail in the previous section. We showed that the absence of high-frequency roll-off in \( F(j\omega) \) demonstrates the roll-off observed in experimental data. In this section we consider a beam model that has both fluid and internal damping. A model of the beam with fluid and internal (Voigt) damping displays high-frequency roll-off commonly observed in experiments.

It is to be noted that a beam model with both fluid and internal damping has been considered by Scherer et al [13]. But the authors of [13] do not argue that internal damping is necessary to yield a physically meaningful model. Furthermore, since in [13] the base is assumed to be clamped, the behavior of the transfer function \( F \) can not be observed. In fact, since Sader’s fluid damping model predicts an inaccurate peak of the first resonant frequency, Scherer et al include internal damping only to match the first resonant peak of the magnitude plot of \( G \) with that obtained from measurements.

Recall that the effect of fluid damping is captured by setting
\[ \tilde{f}(j\omega) = \kappa \omega^2 \Gamma(\omega)^\ast \]
in (3)–(4), and the effect of internal damping is captured by the Voigt damping factor \( \alpha \). The equations of motion describing the beam can thus be written as
\[
\begin{align*}
\psi_{xxxx} &= \frac{\omega^2}{EI} (\mu + \frac{\beta}{j\omega}) \psi, \quad 0 < x < l, \\
\psi(0, j\omega) &= w(j\omega), \\
\psi_x(0, j\omega) &= \psi_x(l, j\omega) = 0, \\
\psi_{xxx}(l, j\omega) &= \frac{1}{EI} \left( \frac{1}{\alpha j\omega + 1} \right) u(j\omega).
\end{align*}
\]

Comparing (10) and (3) it is clear that the transfer functions \( F(j\omega) \) and \( G(j\omega) \) corresponding to system (10) can
be found from (5) and (6) using the following substitution

\[ q(j\omega) = -\frac{\omega^2}{EI} \left( \mu + \kappa \Gamma(\omega)^* \alpha j\omega + 1 \right). \]

Fig. 5 compares the magnitude and phase plots of \( F \) (red) and \( G \) (blue) for Sader’s system (i.e., zero Voigt damping) and a system with both fluid and internal damping as described by (10) with parameter values given in (8) and \( \alpha = 5 \times 10^{-8} \text{s} \). This value of \( \alpha \) is the same as that used in [13]. The plots corresponding to a model with both fluid and Voigt damping are qualitatively similar to those obtained from experimental measurements.

Remark 2: This paper is clearly not the first to consider base oscillations of a beam [14]–[16] or to use the Voigt damping model [13]. Furthermore, it is well-known that the beam equation with viscous damping has an infinite number of poles parallel to the imaginary axis [11] whereas the equation with Voigt damping has poles that bend into the left-half of the complex plane [11] and corresponds to an analytic semigroup [12]. However, to the best of our knowledge, the existing literature on the beam problem does not explicitly point out the shortcomings – with regards to high frequency behavior – of the widely-used viscous-damped and fluid-damped beam models. We believe that the following factors are responsible for this.

(i) High frequency phenomena will not be correctly observed if a Galerkin projection of the governing PDE is made onto a finite set of basis functions [14]; only an infinite-dimensional model is capable of capturing the true high frequency behavior.

(ii) If a spatial impulse is used to imitate an input at the base of the beam [11] the resulting transfer function will be different from that found by considering a genuine base input (i.e., an input that enters the PDE as a time-dependent boundary condition). This fact is acknowledged in the mathematics literature (see for example [17, Example 3.3.1]) but does not seem to have been recognized in the context of the beam problem. And it can be shown that only transfer functions that are obtained from genuine boundary inputs are capable of displaying the correct high frequency response.

(iii) Even if one avoids the pitfalls described above, it is possible to miss interesting high frequency phenomenon if one does not compute the base-to-tip transfer function \( F(s) \). Indeed, as we have shown in this paper, the tip-to-tip transfer function \( G(s) \) displays physically acceptable high frequency roll-off.

V. CONCLUSIONS

We consider the cantilever beam of an AFM operating in dynamic mode. We use the transfer function of base-to-tip displacement to point out important shortcomings in a widely used model for the cantilever beam described in [1]. Since dissipation is the dominant mechanism at high temporal frequencies, it is expected that the mentioned transfer function shows high-frequency roll-off. We demonstrate that this is indeed not the case for the cantilever model proposed in [1]. We show that realistic roll-off behavior is obtained...
when Voigt internal (i.e., structural) damping is added to the model.

Future work in this direction would include identification of the Voigt damping parameter using the transfer function of base-to-tip displacement. The multiplicity expansions proposed in [18] will be utilized in this effort to obtain a rational approximation of the base-to-tip transfer function.

APPENDIX

The $s$-parameterized system (3) represent an ordinary differential equation (in $z$), which admits the following two-point-boundary-value state-space realization

\[
\begin{align*}
\frac{\partial}{\partial x} \Psi &= A_s \Psi, \quad (11a) \\
\Psi &= C \Psi, \quad (11b) \\
B_s v(s) &= N_0 \Psi(0, s) + N_1 \Psi(l, s), \quad (11c)
\end{align*}
\]

with

\[
\begin{align*}
\Psi &= \begin{bmatrix} \psi \\ \psi_x \\ \psi_{xx} \\ \psi_{xxx} \end{bmatrix}, \quad A_s = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -q(s) & 0 & 0 & 0 \end{bmatrix}, \\
C &= \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \quad q(s) = \frac{\mu s^2 - \bar{f}(s)}{EI(\alpha s + 1)}, \\
N_0 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad N_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\
B_s &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} w \\ u \end{bmatrix}.
\end{align*}
\]

We are interested in finding the output $z$ at point $x = l$, i.e., we seek

\[y(s) = z(l, s) = C \Psi(l, s).\]

But $\Psi(l, s)$ is not fully known; equation (11c) only specifies $\psi_3(l, s) = 0$, $\psi_4(l, s) = u(s)$ and does not provide information about the values of $\psi_1(l, s)$ and $\psi_2(l, s)$. To find $\Psi(l, s)$ we first note that

\[\Psi(x, s) = e^{A_s(x - l)} \Psi(l, s),\]

and thus

\[\Psi(0, s) = e^{-A_s l} \Psi(l, s). \quad (12)\]

Substituting (12) into (11c) we obtain

\[(N_0 e^{-A_s l} + N_1) \Psi(l, s) = B_s v(s) \quad (13)\]

which gives

\[\Psi(l, s) = (N_0 e^{-A_s l} + N_1)^{-1} B_s v(s) = e^{A_s l} (N_0 + N_1 e^{A_s l})^{-1} B_s v(s).\]

Note that since $v$ is the only input to the system and all other boundary conditions are homogeneous, then if $v$ is identically equal to zero we would expect $\Psi \equiv 0$ to be the unique solution, and in particular, $\Psi(l, s) \equiv 0$. Now if $v \equiv 0$ then from (13) we have $(N_0 e^{-A_s l} + N_1) \Psi(l, s) \equiv 0$.

Thus if $\Psi(l, s) \equiv 0$ is to be the unique solution we need that $\det(N_0 e^{-A_s l} + N_1) \neq 0$, for all $s$. Therefore the existence of the inverse $(N_0 e^{-A_s l} + N_1)^{-1}$ for all $s$ is required in order to guarantee well-posedness of system (11).

Using $y(s) = C \Psi(l, s)$, we arrive at

\[y(s) = C e^{A_s l} (N_0 + N_1 e^{A_s l})^{-1} B_s v(s) = \begin{bmatrix} F(s) & G(s) \end{bmatrix} \begin{bmatrix} w(s) \\ u(s) \end{bmatrix}, \]

where Mathematica is employed to obtain the transfer functions $F(s)$ and $G(s)$ given in (5) and (6).

REFERENCES