Numerical Methods for a Nonlinear Impact Model: A Comparative Study With Closed-Form Corrections

Stefano Papetti, Federico Avanzini, and Davide Rocchesso

***Abstract*—A physically based impact model—already known and exploited in the field of sound synthesis—is studied using both analytical tools and numerical simulations. It is shown that the Hamiltonian of a physical system composed of a mass impacting on a wall can be expressed analytically as a function of the mass velocity during contact. Moreover, an efficient and accurate approximation for the mass outbound velocity is presented, which allows to estimate the Hamiltonian at the end of the contact. Analytical results are then compared to numerical simulations obtained by discretizing the system with several numerical methods. It is shown that, for some regions of the parameter space, the trajectories of the discretized systems may significantly drift from the analytically derived curves. Two approaches, based on enforcing numerical energy consistency, are then proposed to improve the accuracy of numerical simulations.**

***Index Terms*—Energy conservation, impact modeling, nonlinear dynamical systems, numerical simulation, physics computing.**

I. Introduction

PHYSICAL models of impacts between objects are ubiquitous in many areas of science and engineering, including robotics [1], haptics [2], computer graphics [3], acoustics [4] and sound synthesis [5]. The phenomenologically plausible and energy-consistent behavior of contacting bodies is especially crucial in simulations of interactions based on sustained or repeated impacts, such as in rolling [6], scraping, or bouncing [7].

The higher the upper limit of the perceptual bandwidth is (and the rendering rate), the more critical the accuracy that real-time numerical simulations can afford; thus, making the problem of impact modeling increasingly complex when moving from graphic, to haptic, to auditory displays. Whereas for most graphic displays it is sufficient to describe an impact in terms of the ratio between outbound and inbound velocities, in haptic display and in sound synthesis the perceived characteristics of the impact depend on how bodies interact during contact. Therefore, more sophisticated impact models and carefully designed discretizations are necessary in audio and haptic contexts.

Manuscript received November 10, 2009; revised June 25, 2010 and November 12, 2010; accepted February 03, 2011. Date of publication February 22, 2011; date of current version July 22, 2011. This work was supported by the European Community's Seventh Framework Program under FET-Open grant agreement 222107 NIW-Natural Interactive Walking. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Richard C. Rose.

S. Papetti is with the Computer Science Department, University of Verona, 37134 Verona, Italy (e-mail: stefano.papetti@univr.it).

F. Avanzini is with the Department of Information Engineering, University of Padova, 35131 Padova, Italy (e-mail: avanzini@dei.unipd.it).

D. Rocchesso is with the Department of Art and Industrial Design, IUAV University of Venice, 30135 Venice, Italy (e-mail: roc@iuav.it).

Digital Object Identifier 10.1109/TASL.2011.2118204

A. Impact Models

The classic starting point is the Hertz model of collision between two spheres, which can be extended to include internal viscosity [8]. The impact force in such models is the sum of a nonlinear elastic term—in the form of a power law of compression—and a dissipative component proportional to the compression velocity—via a second power law of compression. For example, see [9] for the case of two colliding spheres.

In the context of musical acoustics, Stulov proposed a piano hammer model that includes relaxation properties of felt [10]. Other models exist that take plastic deformations into account, thus introducing abrupt direction changes in the force-compression curves at the transition between loading and unloading [11].

Particularly popular is the model by Hunt and Crossley [1], [12]–[15], that generalizes the extended Hertz model by considering a variable exponent that accounts for different contact shapes. In this model—described in detail in Section II—the power laws in the elastic and dissipative term are considered to be equal, thus allowing easier closed-form calculations [16]. Despite not being fully justified in physical terms, the Hunt–Crossley model has been quite successful in some areas of engineering because it allows to derive the phase trajectories in closed form, and because it is sufficiently complex to represent a wide variety of contact phenomena. In our work, we adopted this model and extended the range of the available analytical results.

B. Applications in Acoustic Modeling

Contact models can serve as a basis for developing models of acoustic phenomena. In the context of physically based sound synthesis, the Hunt–Crossley model has been used to develop an impact sound model [5], where a generic resonating object is used in place of a rigid wall.

Other models of more complex acoustic phenomena have been developed based on the very same impact model studied here. As an example, a *bouncing* sound model [7] has been obtained by superimposing a constant force, which simulates gravity, on a plain impact sound model. Also, a *rolling* sound model [6] has been implemented by driving an impact sound model by means of a physically inspired control layer. More precisely, the continuous interaction of a ball rolling on a surface has been modeled as a dense temporal sequence of micro-impacts driven by the geometry of the contacting surfaces, and modulated by the ball's asymmetry.

Accurate and consistent impact modeling is crucial in some audio-haptic rendering applications, such as interactive floors or shoes with vibratory and sonic augmentation [17].

In the context of musical sound synthesis, the piano and other percussive musical instruments have also been modeled by using dissipative impact models [10].

C. Issues With Discrete Time

A wide range of numerical methods can be employed to discretize the interaction of impacting bodies. Given a reference continuous-time system, such as the Hunt–Crossley model, the goal is to obtain numerical quantities that follow the continuous-time trajectories as closely as possible, at an affordable computational cost. Thus, efficiency and accuracy are central issues.

Another important goal is energy consistency, especially in the case of repeated or sustained contacts. A numerical method, albeit being accurate, can introduce spurious oscillations or instabilities if it fails in terms of energy conservation. In this regard, the use of wave-digital structures (e.g., wave digital filters [18]) allows to preserve the passivity of certain continuous-time systems. Another approach is that of energy-based methods, i.e., numerical schemes based on the definition of a numerical energy that is proved to be conserved in lossless conditions [19]. These provably stable discrete-time models can be derived for some continuous-time models, including some nonlinear oscillators [20], but are not available for more general nonlinear contact models.

As for applications which make use of impact/contact models, energy inconsistencies are a recurring issue. In computer graphics, where the constraint of low frame rates makes numerical systems prone to instabilities [3], a typical example is provided by a steady object in resting contact with a rigid floor: when the system does not retain passivity, the object can move upward and bounce [21]. Similar issues are encountered in simulations of haptic contact, where stiffness values are usually limited by requirements on system passivity [22], [23], whereas higher values can cause the system to become unstable, for example oscillating, or reacting actively to the input. In numerical sound synthesis by physical models [24], artifacts and inconsistencies can become audible especially in situations of sustained or repeated contact interactions, as in rolling, sliding, or bouncing.

Instead of aiming at provably stable numerical methods, this study looks at the accuracy of some methods commonly found in physics-based engines [3], and measures their performance in following the theoretical phase trajectories and in reproducing the energy exchanges that occur in the continuous-time impact model.

D. Outline

In this paper, the Hunt–Crossley impact model is first characterized by proving some novel analytical results. Such results are then used as a reference to compare the accuracy of several numerical simulations of the model, obtained by discretizing the continuous-time equations with a number of widely used numerical methods. It is shown that for some regions of the parameter space, the trajectories of the discretized systems may significantly drift from the analytically derived curves. Finally, by exploiting the provided analytical results, two approaches are proposed which improve the accuracy of the numerical simulations, thus restoring their energy consistency.

More in detail, the main contribution of this work can be summarized as follows.

1) Novel approximate closed-form expression for the outbound velocity, and quantification of its relative error. Expression of the total energy (Hamiltonian) as a function of compression velocity (Section II-A).
2) Analysis of the distortion caused by a constant external force on the analytically derived phase portraits. This is of some importance for applying the results to real-world simulations (Section II-B).
3) Comparison of the analytically derived phase portrait and Hamiltonian with those obtained by applying four relevant numerical methods to the Hunt–Crossley impact model. Two critical cases are examined: low dissipation and hard impact (Section III-B).
4) Proposal of two correction methods for numerical models, one based on the analytical dependence of compression on velocity, and one based on a constraint on the outbound velocity (Section IV).

The remainder of the paper is organized as follows. Section II contains the analytical study, Section III-A describes different numerical methods which are used for discretizing the continuous-time system, while in Section III-B the corresponding numerical simulations are compared. Sections IV-A and IV-B show how the provided analytical results can be used to improve the behavior of the numerical simulations. Finally, in Sections IV-C and IV-D the computational cost and accuracy of simulations with and without corrections are compared and evaluated.

II. Impact Model

The Hunt–Crossley impact model [12] is described by the following nonlinear equation for the impact force:

$$f(x,v) = \begin{cases} kx^{\alpha} + \lambda x^{\alpha} v = kx^{\alpha} \cdot (1 + \mu v), & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (1)$$

where x is the *compression*, $v = \dot{x}$ is the *compression velocity*, $\alpha > 1$ is the exponent of a power law and represents the *local shape* of contact surfaces, k is the *stiffness coefficient*, and $0 \leq \lambda \leq k$ is the *damping coefficient*. The mathematically convenient term $\mu\ (= \lambda/k)$ allows to simplify some closed-form calculations. The impact force model thus represents a nonlinear spring of constant k in parallel with a nonlinear damper of constant λ. The term kx^{α} corresponds to the *elastic component*, while $\lambda x^{\alpha} v$ represents the *dissipation* due to internal friction.

Marhefka and Orin [1] made use of the Hunt–Crossley model in order to represent the impact between a single mass m and a comparatively massive object which does not move during collision, therefore considering the system described by the equation

$$ma(t) = -f\left(x(t), v(t)\right) \quad (2)$$

where a is the *acceleration* of the mass m.

In order to simplify the study of the model and avoid details on the geometry of the contact surfaces, one can equivalently consider a point-mass impacting on a wall, where the contact is represented by the Hunt–Crossley model [13]. According to this

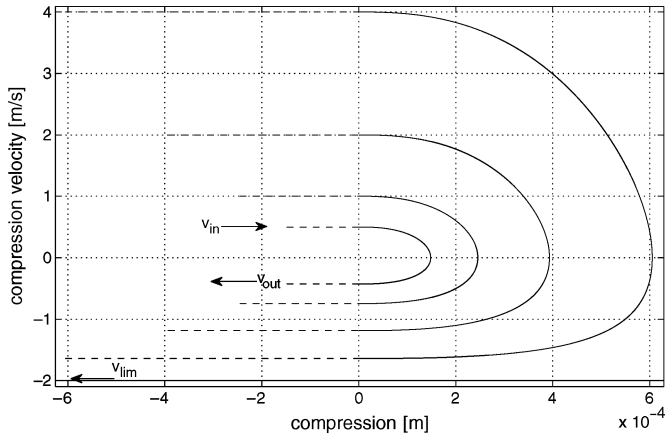


Fig. 1. Phase portraits for varying input velocities: $v_{\rm in}$ =0.5, 1, 2, 4 m/s. Other values of parameters are: $m = 10^{-2}$ Kg, $k = 10^{7}$ N/m$^{\alpha}$, $\mu = 0.5$ s/m, $\alpha = 1.5$. Solid lines represent the mass trajectory during contact; dashed lines represent free motion.

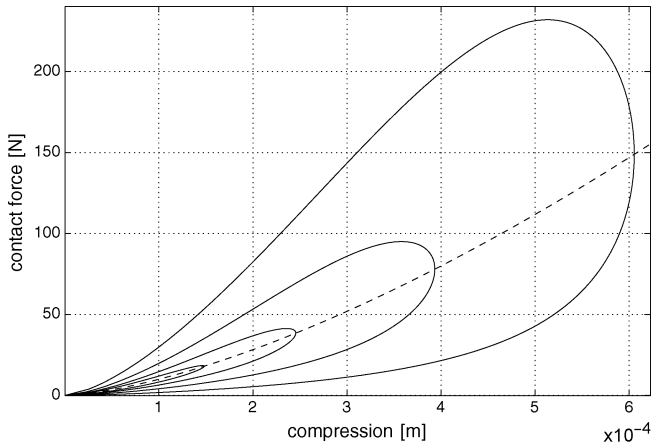


Fig. 2. Compression-force characteristics for varying input velocities: $v_{\rm in}$ =0.5, 1, 2, 4 m/s. Solid lines represent the case when dissipation is taken into account (the values of parameters are the same as in Fig. 1). The dashed line represents the curve kx^{α}, where no dissipation is considered ($\lambda = 0$).

configuration—represented by the system of (2) and (1)—the compression x and the compression velocity v are respectively defined as the position and velocity of the point-mass with respect to the wall (zero reference). The position and velocity are considered positive, respectively, while the mass is in contact with the wall, and when the mass is directed towards the latter.

A. Properties and Analytical Results

Thanks to the simple form of (2), the model can be treated analytically and some of its properties can be inferred. Hereafter, the initial conditions $x(0) = 0$ and $\dot{x}(0) = v_{\rm in}$ are considered, that is to say that the mass hits the wall with velocity $v_{\rm in}$ at time $t = 0$.

1) Compression: It is shown in [1] that from (2) it follows:

$$x(v) = \left[\frac{m(\alpha+1)}{k\mu^2}\cdot\left(-\mu(v-v_{\rm in}) + \log\left|\frac{1+\mu v}{1+\mu v_{\rm in}}\right|\right)\right]^{\frac{1}{\alpha+1}} \tag{3}$$

which can be exploited for plotting the phase portraits on the (x, v) plane shown in Fig. 1. From Fig. 1 it can be inferred that, due to the viscous dissipation occurring during contact, the relation $v(t+dt) < v(t)$ holds, and in particular the output velocity $v_{\rm out}$ is always smaller in magnitude than the corresponding $v_{\rm in}$. Moreover, for increasing $v_{\rm in}$'s, $v_{\rm out}$ converges to the limit value $v_{\rm lim} \triangleq -1/\mu$. The line $v = v_{\rm lim}$ represents the trajectory where the elastic and dissipative terms cancel, and separates two regions of the phase space, each of which is never entered by trajectories started in the other one.

Equation (3) allows to infer the *maximum compression* experienced during contact, which occurs when the compression velocity equals zero:

$$x_{\max} = x(0) = \left[\frac{m(\alpha+1)}{k\mu^2}\cdot\left(\mu v_{\rm in} + \log\left|\frac{1}{1+\mu v_{\rm in}}\right|\right)\right]^{\frac{1}{\alpha+1}}. \tag{4}$$

Finally, by substituting (3) in (2) one can plot the compression-force characteristics during collision, which are shown in Fig. 2. It can be noted that the dissipative term $\lambda x^{\alpha} v$ introduces hysteresis around the curve kx^{α}.

2) Output Velocity: The *restitution coefficient* E is defined as

$$E \triangleq \left|\frac{v_{\rm out}}{v_{\rm in}}\right|. \tag{5}$$

Note that $v_{\rm in}$ and $v_{\rm out}$ correspond to the roots of the right-hand side of (3), that is the points where $x = 0$. As a result, $v_{\rm out}$ can be defined implicitly from (3) as a function of $(\mu, v_{\rm in})$ only:

$$\mu v_{\rm out} - \log|1+\mu v_{\rm out}| = \mu v_{\rm in} - \log|1+\mu v_{\rm in}|. \tag{6}$$

This implies that $\mu v_{\rm out}$ is a function of $\mu v_{\rm in}$ only, and therefore E is also a function of $\mu v_{\rm in}$ only.

Analytical derivations of the dependence $E(\mu v_{\rm in})$ have been classically performed in the limit of small initial velocities and/or small dissipation [12].[1] However, we suggest that a nonlocal approximation $\tilde{v}_{\rm out}$ can be empirically determined as an *ansatz* which fits the curve $E(\mu v_{\rm in})$ in the two limit regions $\mu v_{\rm in} \to 0_+$ and $\mu v_{\rm in} \to +\infty$, thus obtaining:[2]

$$\tilde{v}_{\rm out}(\mu, v_{\rm in}) = v_{\rm lim}\left[1 - \left(\sum_{j=0}^{n} b_j \cdot v_{\rm in}^{j}\right) e^{-2\mu v_{\rm in}}\right] \tag{7}$$

where, in the case $n = 4$, the coefficients b_j are

$$b_0 = 1, \quad b_1 = \mu, \quad b_2 = \frac{2}{3}\mu^2, \quad b_3 = \frac{2}{9}\mu^3, \quad b_4 = \frac{14}{135}\mu^4. \tag{8}$$

From now on, unless specified otherwise, the notation $\tilde{v}_{\rm out}$ will refer to the fourth-order approximation provided by (7) and the coefficients (8). Conventional iterative zero-finding methods applied to (6) can always be used to compute a more precise release velocity at a higher computational cost (see Section IV-C).

Fig. 3 shows the error introduced by the approximate value $\tilde{v}_{\rm out}$, when compared to the corresponding value computed numerically as a zero of (6).

[1]Appendix A provides an example.

[2]See Appendix B for details.

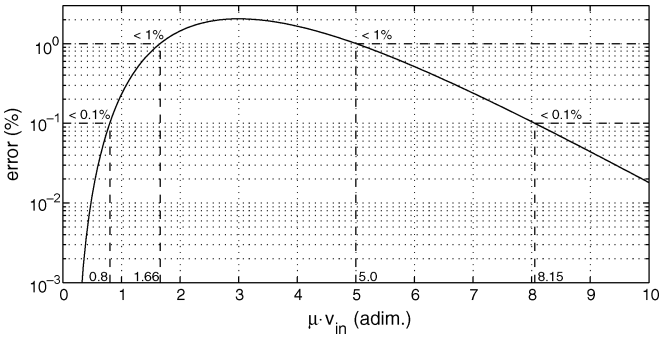


Fig. 3. Log-scaled percentage error of the output velocity approximated by $\bar{v}_{\rm out}$, with respect to the value computed numerically as a zero of (6). Ranges of $\mu v_{\rm in}$ are shown, for which the maximum error is respectively less than 1% and 0.1%.

3) Contact Time: It is shown in [25] that the *contact time* can be expressed as

$$\tau = \left(\frac{m}{k}\right)^{\frac{1}{\alpha+1}} \cdot \left(\frac{\mu^2}{\alpha+1}\right)^{\frac{\alpha}{\alpha+1}} \cdot \int_{v_{\rm out}}^{v_{\rm in}} \frac{1}{(1+\mu v)\left[-\mu(v-v_{\rm in}) + \log\left|\frac{1+\mu v}{1+\mu v_{\rm in}}\right|\right]^{\frac{\alpha}{\alpha+1}}}. \quad (9)$$

Equation (9) states that the contact time τ depends only on μ, the exponent α and the ratio m/k, plus obviously the impact velocity $v_{\rm in}$. Since neither m nor k affect the value of the integral (recall that $v_{\rm out}$ depends only on μ and $v_{\rm in}$), it follows that, given a fixed $v_{\rm in}$, the proportionality $\tau \sim (m/k)^{1/(\alpha+1)}$ holds.

From an auditory point of view the value of the contact time is strongly correlated to the perceived "hardness" of the impact [25], [26]. Namely, as the contact time decreases, the perceived hardness increases. Recalling the power-law dependence above and (1) it follows that, for a fixed mass m, "hard" and "soft" impacts correspond respectively to high and low force values.

4) Energy Properties and Behavior: The energy variation in a mechanical system can be calculated as the work made by the non-conservative forces $f_{\rm nc}$ acting on the system along a certain path $x_1 \rightarrow x_2$

$$\Delta H = \int_{x_1}^{x_2} f_{\rm nc}(x)dx = \int_{t_1}^{t_2} f_{\rm nc}(t)v(t)dt = -\Delta\Lambda \quad (10)$$

where H is the total energy content, known as the *Hamiltonian*, Λ is the energy dissipation, and the second integral is obtained by application of the chain rule ($dx(t) = v(t)dt$) and considering that t_1 and t_2 correspond respectively to the instants when the displacements x_1 and x_2 are reached. The Hamiltonian H is the sum of potential and kinetic energies, hereafter named V and T, respectively,

$$H(t) = V(t) + T(t). \quad (11)$$

With regard to the system represented by (2), T is related to the dynamics of the mass, which is described by the left-hand side of (2), while V is related to the elastic component of the impact force of (1).

In agreement with the last integral in (10), multiplying both sides of (2) by $v(t) = dx/dt$ and time-integrating them, gives

$$\underbrace{\int_0^t ma(s)v(s)ds}_{T(t)-T_0} = -\underbrace{\int_0^t kx(s)^\alpha v(s)ds}_{V(t)} - \underbrace{\int_0^t \lambda x(s)^\alpha v(s)^2 ds}_{\Lambda(t)>0} \quad (12)$$

where the force expression of (1) has been considered in the case $x > 0$ only. The first two integrals in (12) can be solved explicitly, obtaining

$$V(t) = \frac{kx(t)^{\alpha+1}}{\alpha+1}, \quad T(t) = \frac{mv(t)^2}{2}. \quad (13)$$

Considering a system where the mass travels with velocity $v_{\rm in}$ before an impact occurs, then the *initial Hamiltonian* corresponds to the initial kinetic energy

$$H_0 = T_0 = \frac{mv_{\rm in}^2}{2}. \quad (14)$$

From (10) it follows that at each time instant t:

$$H(t) = H_0 - \Lambda(t) \quad (15)$$

and, since $\Lambda(t) > 0$, the following inequalities hold:

$$0 \le H(t+dt) \le H(t). \quad (16)$$

Indicating τ as the instant when an impact ends, the *final Hamiltonian* of the system, that is the energy content right after contact, can be written as

$$H_\tau = T_\tau = \frac{mv_{\rm out}^2}{2}. \quad (17)$$

Also, the total amount of energy dissipation occurred during contact is

$$\Delta H_\tau = H_\tau - H_0 = -\Lambda_\tau. \quad (18)$$

Since the area enclosed by the hysteresis loops shown in Fig. 2 is by definition the work done by the impact force, such area corresponds to ΔH_τ.

As for the rightmost integral in (12), which is non-solvable, an equivalent expression can be obtained by substituting the complementary results for the remaining integrals:

$$\Lambda(t) = \int_0^t \lambda x(s)^\alpha v(s)^2 ds = \frac{m\left(v_{\rm in}^2 - v(t)^2\right)}{2} - \frac{kx(t)^{\alpha+1}}{\alpha+1}. \quad (19)$$

Finally, by substituting (3) in (19) and recalling (15), the following expression in v only is found:

$$H(v) = H_0 - \Lambda(v) = \frac{m}{2}v^2 - \frac{m}{\mu}(v - v_{\rm in}) + \frac{m}{\mu^2}\log\left|\frac{1+\mu v}{1+\mu v_{\rm in}}\right| \quad (20)$$

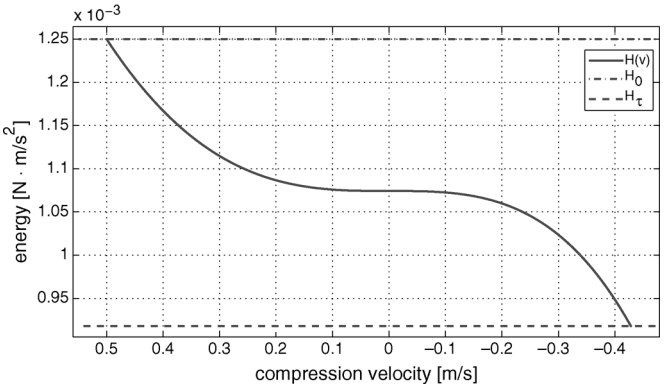


Fig. 4. Compression velocity–Hamiltonian characteristic. The two horizontal lines display $H_0 = T_0$ and $H_\tau = T_\tau$, that is, respectively, the initial and the final Hamiltonian. The values of parameters are the same as in Fig. 1 for $v_{\rm in} = 0.5$ m/s. The compression velocity axis has been inverted, thus allowing to read the graph from left to right.

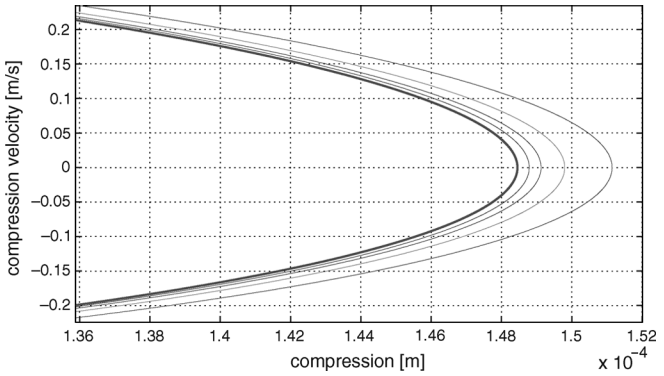


Fig. 5. Detail of phase portraits for different external forces applied: $f_e = m \cdot g \cdot (0.5, 1, 2, 4)$ N, where $g = 9.81$ m/s^2 is the standard gravity. The bold line represents the case where $f_e = 0$. The values of parameters are the same as in Fig. 1 for $v_{\rm in} = 0.5$ m/s. Notice that, due to the lack of analytical results when an external force is present, the phase portraits have been obtained as the result of numerical simulations.

which can be used for plotting the curve shown in Fig. 4.

B. Addition of a Constant External Force

When a constant external force f_e (e.g., the force of gravity) is applied to the mass, (2) has to be rewritten as

$$ma(t) + f_e = -f\left(x(t), v(t)\right). \tag{21}$$

Unfortunately, in this case no closed-form analytical results can be found as those described in Section II-A. In more detail, multiplying both sides of (21) by $v(t)$ and time-integrating them, an equation is found where an unsolvable integral is present, this way preventing to directly obtain explicit-form expressions for $x(v)$, $\tilde{v}_{\rm out}$ and τ.

Rewriting the second-order equation in (21) as a system of first-order equations

$$\begin{cases} \dot{x} = v \\ \dot{v} = -\frac{k}{m}x^\alpha - \frac{\lambda}{m}x^\alpha v - \frac{f_e}{m} \end{cases} \tag{22}$$

the equilibrium point of the system is found to be $(x_{\rm eq}, v_{\rm eq}) = (-f_e^{1/\alpha}/k, 0)$, which corresponds to the *compression offset* in stationary conditions.

As Fig. 5 shows, for positive values of f_e and $v_{\rm in}$, the velocity of the mass during the compression phase is generally greater than in the case when $f_e = 0$. In particular, at the beginning of contact interaction, since f_e is higher than the current impact force f, the compression velocity exceeds $v_{\rm in}$. On the other hand, compared to the case when $f_e = 0$, during the decompression phase the absolute value of the mass velocity v decreases, resulting in lower output velocities. Moreover, the resulting maximum compression is always greater than that calculated by (4).

III. Numerical Simulations

In this section, the continuous-time system described by (2) is discretized by means of several numerical methods, and the resulting numerical systems are studied.

A. Numerical Methods

Different numerical methods were considered, which are commonly used in various fields of applications ranging from computer graphics, to physical simulation of dynamical systems and digital signal processing. Thanks to their low order—which generally results in low computational costs—the chosen methods are particularly suitable for real-time applications:

- the trapezoid rule is popular to translate analog filter structures to discrete-time filters, and it is the basis for wave digital filters;
- Verlet integration is popular in physics-based graphic engines;
- Heun is a second-order method (complexity similar to the previous two) representative of the Runge–Kutta family;
- fourth-order Runge–Kutta is expected to be more accurate, and more expensive.

Following the standard notation in numerical analysis, the integration step is a constant named h $(= 1/F_s)$.

1) 1-Step Adams-Moulton (AM1): Is a A-stable second-order implicit method [27], also known as *bilinear transformation*, or *trapezoid rule*.

Discretizing (2) results in the following equation in state-space form:

$$\begin{bmatrix} x_{n+1} \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_n \\ v_n \end{bmatrix} + \begin{bmatrix} \frac{h^2}{4m} \\ \frac{h}{2m} \end{bmatrix} [f_{n+1} + f_n] \tag{23}$$

where the expression for the discrete-time force is obtained by replacing the continuous-time variables $x(t)$ and $v(t)$ in (1) with their discrete-time counterparts.

Since the AM1 method is implicit, (23) is also in implicit form, and this is reflected in the instantaneous relationship between $[x_{n+1}\ v_{n+1}]^T$ and f_{n+1}. Unfortunately, since f_{n+1} also has an instantaneous dependence on x_{n+1} and v_{n+1} given by (1), the discrete-time counterpart of the system described by (2) contains a *delay-free loop*, which is not directly computable and—because of the nonlinear dependence $f(x, v)$—needs some special handling in order to be solved. In particular, the *K-method* [28] together with *Newton's method* [27] are used, weighing on the efficiency of the simulation (see Section IV-C).

2) Verlet: Is a second-order explicit method commonly used in computer graphics [3], video games, and molecular dynamics simulation, where it is typically used for integrating Newton's equation of motion in order to describe the trajectory of moving

particles. The one used here is a variant, called *velocity Verlet*, which provides better handling of the velocity variable and can be seen as a predictor-corrector method.

Discretizing (2) results in the following implementation scheme:

$$\begin{aligned} x_{n+1} &= x_n + hv_n + \frac{h^2}{2}\frac{f_n}{m} \\ v_{n+\frac{1}{2}} &= v_n + \frac{h}{2}\frac{f_n}{m}, \quad \text{predictor} \\ f_{n+1} &= f\left(x_{n+1}, v_{n+\frac{1}{2}}\right) \\ v_{n+1} &= v_{n+\frac{1}{2}} + \frac{h}{2}\frac{f_{n+1}}{m}, \quad \text{corrector.} \end{aligned} \tag{24}$$

It should be noted that this algorithm assumes that f_{n+1} only depends on the predicted velocity $v_{n+(1/2)}$, which clearly gives rise to inaccuracies.

3) Heun: Is a predictor-corrector explicit method [3], [27], with the forward Euler method as predictor and the trapezoid rule as corrector. It can also be seen as a *second-order Runge–Kutta method* (RK2).

Discretizing (2) results in the following implementation scheme:

$$\begin{aligned} \tilde{v}_{n+1} &= v_n + h\frac{f_n}{m}, \quad \text{predictor} \\ x_{n+1} &= x_n + \frac{h}{2}(v_n + \tilde{v}_{n+1}) \\ f_{n+1} &= f(x_{n+1}, \tilde{v}_{n+1}) \\ v_{n+1} &= v_n + \frac{h}{2}\frac{f_n + f_{n+1}}{m}, \quad \text{corrector.} \end{aligned} \tag{25}$$

Again, it should be noted that both x_{n+1} and f_{n+1} only depend on the predicted velocity $\tilde{v}_{n+1}$, and this gives rise to inaccuracies.

4) Fourth-Order Runge–Kutta (RK4): Is an explicit iterative method [3], [27] which is widely used to solve ODEs with improved accuracy.

Discretizing (2) results in the following implementation scheme:

$$\begin{aligned} x_{n+1} &= x_n + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4) \\ v_{n+1} &= v_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned} \tag{26a}$$

where

$$\begin{aligned} l_1 &= hv_n, \quad l_2 = h\left(v_n + \frac{k_1}{2}\right) \\ l_3 &= h\left(v_n + \frac{k_2}{2}\right), \quad l_4 = h(v_n + k_3) \\ k_1 &= h\frac{f_n}{m}, \quad k_2 = h\frac{f\left(x_n + \frac{l_1}{2}, v_n + \frac{k_1}{2}\right)}{m} \\ k_3 &= h\frac{f\left(x_n + \frac{l_2}{2}, v_n + \frac{k_2}{2}\right)}{m} \\ k_4 &= h\frac{f(x_n + l_3, v_n + k_3)}{m}. \end{aligned} \tag{26b}$$

It should be noted that, for each sample, both the velocity and the nonlinear force of (1) need to be evaluated four times, therefore strongly affecting the efficiency of the simulation (see Section IV-C).

B. Experimental Results

In order to evaluate the chosen numerical methods, it is useful to compare the behavior of the corresponding simulations against the known analytical results (see Section II-A).

The main references used to quantitatively assess the reliability of a particular numerical method *during* contact are (3) and (20), which express respectively the compression x and the Hamiltonian H as functions of the compression velocity v. The errors in x and H are then measured as the maximum deviations of their discrete-time versions from the respective analytical curves $x(v)$ and $H(v)$, in relation to the entire variation range of the quantities considered (i.e., $x_{\max}$ and ΔH_τ). In detail, taking into account (20) and defining $H^{\text{sim}}(x_n, v_n)$ as the Hamiltonian computed from (13) using compression and velocity values resulting from a numerical simulation,[3] and ΔH_τ as in (18), the *maximum deviation on* H is calculated by means of the following expression:

$$\%\text{dev}(H) = 100 \cdot \left| \frac{\max_n \left\{ H^{\text{sim}}(x_n, v_n) - H(v_n) \right\}}{\Delta H_\tau} \right|. \tag{27}$$

Thus, the deviation is obtained by normalizing the absolute error according to the *range of variation* of the quantities being evaluated, and picking the maximum deviation along such range. It can be viewed as the Chebyshev distance between H^{sim} and H during contact, normalized by ΔH_τ. Such measure has a different meaning compared to ordinary relative errors, since the latter would only account for local deviations from the analytical curves. Similarly to what is described above, considering (3) and defining $x_{\max}$ as in (4), the *maximum deviation on* x is calculated as

$$\%\text{dev}(x) = 100 \cdot \left| \frac{\max_n \left\{ x_n - x(v_n) \right\}}{x_{\max}} \right|. \tag{28}$$

Finally, another indicator which allows to evaluate the accuracy and consistency of the simulations at release from contact is provided by the output velocity computed numerically as a zero of (6), which is used to calculate the relative error $\%\text{err}(v_{\text{out}})$ of its simulated counterpart $v_{\text{out}}^{\text{sim}}$.

Throughout the following example simulations some values of parameters are kept constant: $m = 10^{-2}$ kg, and $F_s = 44.1$ kHz (i.e., a standard audio sample rate).

1) Non-Critical Simulations: In order to verify the numerical implementations, the parameters of the model are set to a "safe" configuration, that is to far-from-extreme values. In this case, contact extends over many samples, thus ensuring that the simulations are influenced only to a negligible extent by the chosen sampling rate (see Section I-C) and should more likely behave as the original continuous-time system.

This is confirmed qualitatively by Fig. 6, where the plots of all such simulations substantially overlap and coincide with the

[3]Not to be confused with the discrete Hamiltonian defined in [24] for finite-difference schemes.

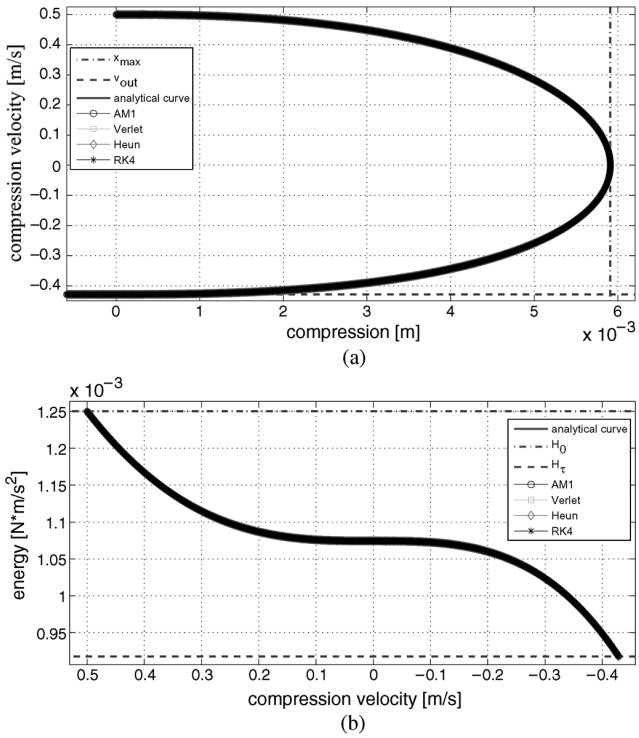


Fig. 6. Qualitative comparison of different methods in noncritical simulations. The values of parameters are $k = 10^3$ N/m$^\alpha$, $\mu = 0.5$ s/m, $\alpha = 1.5$, $v_{\rm in} = 0.5$ m/s. (a) Phase portraits. The two tangent lines represent the maximum compression $x_{\max}$ calculated by (4) and the output velocity $v_{\rm out}$ computed numerically as a zero of (6). (b) Energy behaviors. The two horizontal lines display $H_0 = T_0$ and $H_\tau = T_\tau$, i.e., respectively, the initial and the final Hamiltonian.

TABLE I
SUMMARY OF ERRORS IN NONCRITICAL SIMULATIONS. THE VALUES OF PARAMETERS ARE $k = 10^3$ N/m$^\alpha$, $\mu = 0.5$ s/m, $\alpha = 1.5$, $v_{\rm in} = 0.5$ m/s

errors	AM1	Verlet	Heun	RK4
%dev(x)	0.255	0.018	0.319	0.005
%err($v_{\rm out}$)	$+2\cdot10^{-5}$	$+2\cdot10^{-6}$	$-3\cdot10^{-5}$	$+2\cdot10^{-6}$
%dev(H)	$3\cdot10^{-4}$	0.052	$4\cdot10^{-4}$	$1\cdot10^{-5}$

analytical curves. Moreover, Table I offers a quantitative evaluation of the simulations, showing the errors introduced by the different numerical methods considered.

2) Critical Parameter Regions: It has been found empirically that when $\tau^{\rm sim} \leq 4$ samples, the errors in x, H and $v_{\rm out}$ heavily increase, resulting in an extremely poor reliability of all the simulations. The obvious reason for this behavior is that, since only very few samples of data are available, the numerical systems are totally unable to describe the original continuous-time counterpart. Hence, in the following study, only values of parameters resulting in $\tau^{\rm sim} > 4$ samples are considered.

It has been observed that when $\mu \to 0_+$ and/or as the contact time $\tau^{\rm sim}$ decreases (i.e., for "hard" impacts), the behavior of most numerical implementations tends to become inconsistent with the continuous-time system. Hereafter, the numerical systems are studied for these two critical configurations, respectively, named *case 1* and *case 2*.

Case 1: Low dissipation ($\mu \to 0_+$).

In case of low dissipation, the Hamiltonian of both Verlet- and Heun-discretized systems is prone to oscillations, while

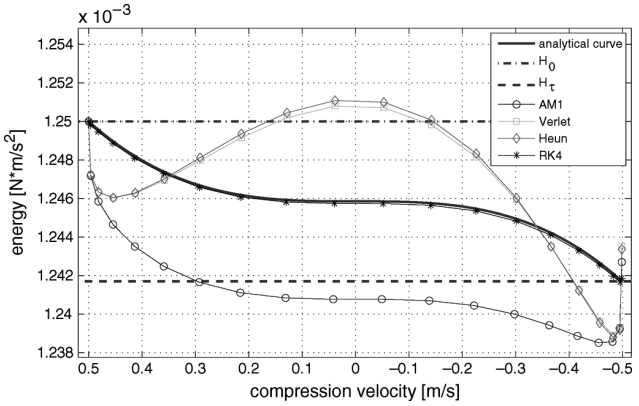


Fig. 7. Comparison of Hamiltonians for different implementations of a simulation example following *case 1*. The values of parameters are $k = 10^7$ N/m$^\alpha$, $\mu = 0.01$ s/m, $\alpha = 1.3$, $v_{\rm in} = 0.5$ m/s. The contact time equals 19 samples.

TABLE II
SUMMARY OF ERRORS IN EXAMPLE SIMULATIONS OF *case 1* AND *case 2*. THE LAST COLUMN SHOWS THE ERROR YIELDED BY THE APPROXIMATE VALUE $\tilde{v}_{\rm out}$ WITH RESPECT TO THE VALUE COMPUTED NUMERICALLY: IT IS WORTH NOTICING THAT IN BOTH CASES THE ERROR IS LOWER THAN THOSE YIELDED BY THE SIMULATIONS. (a) SIMULATIONS FOLLOWING *case 1*. THE VALUES OF PARAMETERS ARE THE SAME AS IN FIG. 7. (b) SIMULATIONS FOLLOWING *case 2*. THE VALUES OF PARAMETERS ARE THE SAME AS IN FIG. 8

(a)

errors	AM1	Verlet	Heun	RK4	$\tilde{v}_{\rm out}$
%dev(x)	1.011	1.083	1.136	0.052	-
%err($v_{\rm out}$)	$+0.039$	$+0.073$	$+0.067$	$+0.006$	$-1\cdot10^{-5}$
%dev(H)	61.302	59.542	63.042	1.427	-

(b)

errors	AM1	Verlet	Heun	RK4	$\tilde{v}_{\rm out}$
%dev(x)	4.381	4.418	19.506	0.412	-
%err($v_{\rm out}$)	$+2.551$	$+0.839$	-4.692	-0.105	-0.013
%dev(H)	7.885	9.475	23.387	0.410	-

contact ends in an inconsistent final energy state: typically, $H_\tau^{\rm Verlet,Heun} > H_\tau$, where H_τ is defined as in (17) and is calculated using values of $v_{\rm out}$ computed numerically as zeros of (6). As for AM1-discretized systems, these generally tend to dissipate too much energy during contact (i.e., $H^{\rm AM1} < H$), while slightly gaining spurious energy as the contact ends (i.e., $H_\tau^{\rm AM1} > H_\tau$). On the other hand, RK4-discretized systems generally behave quite consistently both during and after the contact interaction (i.e., $H^{\rm RK4} \approx H$).

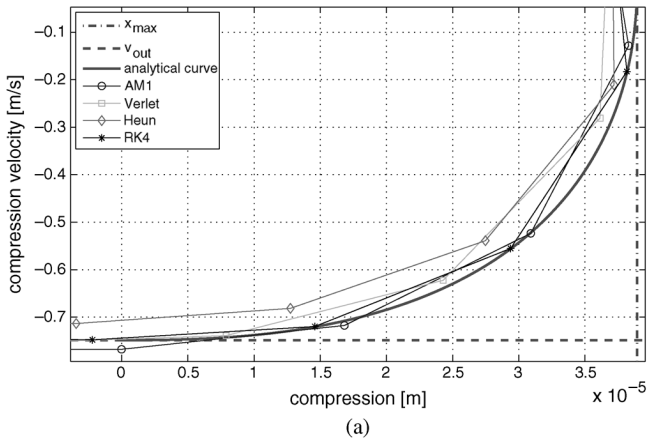
Fig. 7 shows the Hamiltonian of a simulation example with low dissipation ($\mu = 0.01$), while Table II(a) shows the resulting errors.

Case 2: Hard impacts.

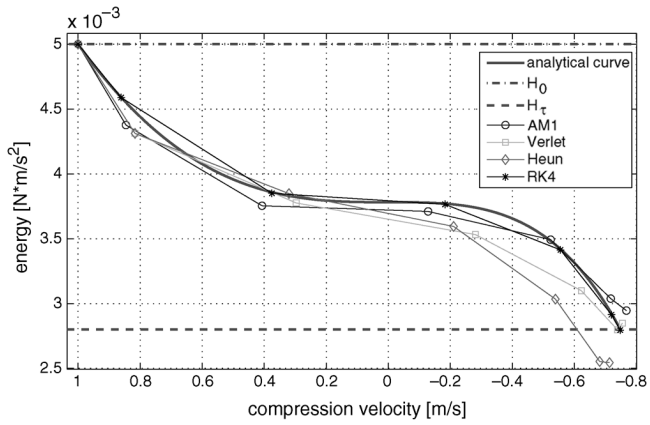
With the exception of RK4-discretized systems, in this case the simulations usually show more spread errors: $x^{\rm sim}$, $H^{\rm sim}$, and $v_{\rm out}^{\rm sim}$ tend to substantially deviate from the respective analytical results.

Fig. 8 shows a hard impact simulation example where the values of parameters are $k = 10^9$ N/m$^\alpha$, $\mu = 0.5$ s/m, $\alpha = 1.5$, $v_{\rm in} = 1$ m/s, while Table II(b) shows the corresponding errors. The resulting contact time τ equals 6 samples.

It is worth noticing that the RK4 method has been proved to behave quite consistently across disparate configurations of



(a)



(b)

Fig. 8. Comparison of phase portraits and Hamiltonians for different implementations of a simulation example following *case 2*. The values of parameters are $k = 10^9$ N/m$^\alpha$, $\mu = 0.5$ s/m, $\alpha = 1.5$, $v_{\rm in} = 1$ m/s. The contact time equals 6 samples. (a) Detail of phase portraits. (b) Hamiltonians.

parameters. Therefore, a highly oversampled RK-discretized system can be taken as a reference, able to provide extremely accurate simulations.

IV. Improved Numerical Simulations

A. Exploitation of Analytical Results

Making use of some of the analytical results obtained in Section II-A, in this section solutions are proposed to fix the inconsistencies pointed out in Section III-B. The aim is to improve the accuracy and reliability of simulations which make use of the impact model under study, in view of their implementation as real-time applications.

1) Hybrid Numerical-Analytical Computation: Instead of computing both the state variables x_n and v_n as the result of a numerical simulation—for example using one of the numerical methods described in Section III-A—this solution consists in computing only the compression velocity v_n *numerically*, then employing it in (3) to calculate the corresponding value of compression $x(v_n)$ *analytically*.

As a result, the corrected numerical system strictly follows the analytical curves $x(v)$ and $H(v)$.

Since the computation is to be made at each sample for the whole duration of the contact interaction, this solution noticeably increases the computational load of the simulation.

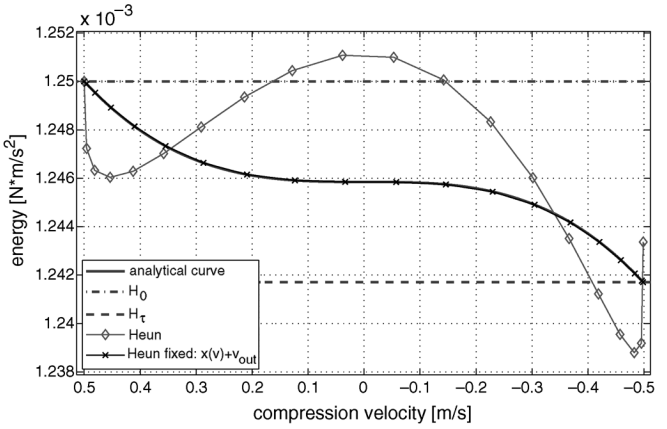


Fig. 9. Comparison of the Hamiltonians of a Heun-based simulation example following *case 1*, with and without corrections. The values of parameters are the same as in Fig. 7.

2) Output Velocity Constraint: The solution above can only be applied while the contact interaction lasts, and therefore the behavior of a numerical system upon the end of the impact is not affected. An additional method is proposed here to control the energy content (i.e., the residual kinetic energy $T_\tau^{\rm sim} = H_\tau^{\rm sim}$) at that time: as soon as x_n takes a non-positive value following an impact (i.e., at the detachment of the two objects), the output velocity can be forced to the approximate value $\tilde{v}_{\rm out}$ given by (7).

Considering that the error introduced by (7) depends solely on the product $\mu v_{\rm in}$ (see Fig. 3), it is advisable to apply the correction only when the product $\mu v_{\rm in}$ corresponds to an acceptable error, or the risk is to even worsen the behavior of the numerical system. However, the use of this conditional correction always implies a tradeoff: on the one hand it guarantees not to introduce errors greater than a chosen maximum, while on the other hand 1) within the excluded range of $\mu v_{\rm in}$, the output velocity is not controlled and therefore depends only on the plain numerical method chosen, 2) within the included intervals of $\mu v_{\rm in}$, the risk is that the correction introduces errors even greater than those provided by the non-corrected numerical system (this is true especially for noncritical parameter regions).

Once the output velocity has been forced to $\tilde{v}_{\rm out}$, the corresponding compression should be set to 0, this way adhering to the compression-force characteristics shown in Fig. 2 by closing their numerical counterpart at $(x = 0, f = 0)$, and ensuring that the final potential energy $V_\tau^{\rm sim}$ is set to zero.

The computation of $\tilde{v}_{\rm out}$ only needs to take place in correspondence to an impact event, and as soon as the impact velocity $v_{\rm in}$ is known.

B. Numerical Simulations With Corrections

In order to test the described corrections, they were applied to the worst behaving simulation examples provided in Section III-B2: Figs. 9 and 10 show a comparison of Heun-based simulations following, respectively, *case 1* and *case 2*, with and without corrections.

1) Improved Energy Behavior: When the hybrid correction described in Section IV-A1 is applied, any simulation strictly adheres to the analytical curves $H(v)$ and $x(v)$ during contact,

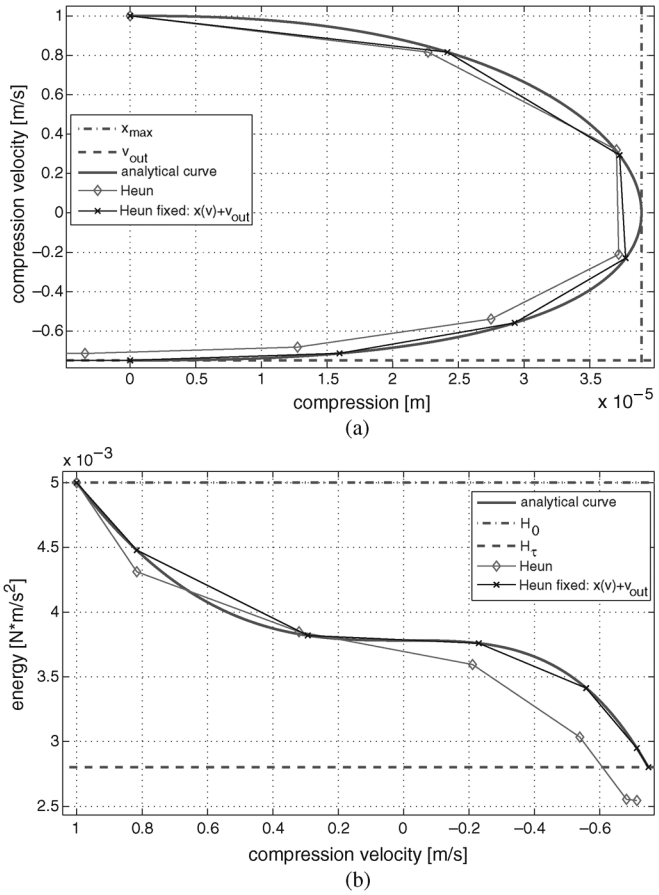


Fig. 10. Comparison of a Heun-based simulation example following *case 2*, with and without corrections. The values of parameters are the same as in Fig. 8. (a) Phase portraits. (b) Hamiltonians.

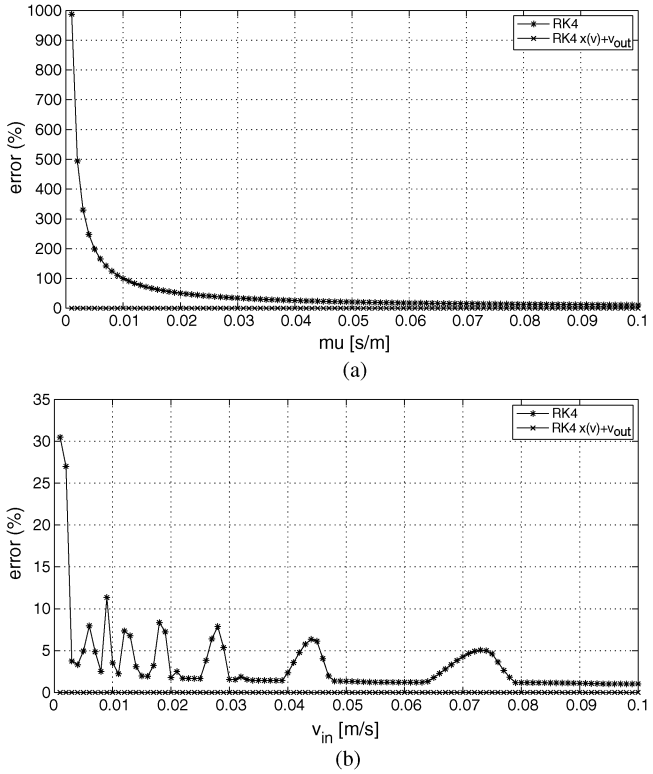


Fig. 11. Trend of error on the Hamiltonian for RK4-based simulations with and without corrections, for small values of μ and $v_{\rm in}$. For the values of $\mu v_{\rm in}$ considered, the output velocity constraint is always applied (see Fig. 3). The values of parameters, where kept constant, are $k = 10^9$ N/m$^\alpha$, $\mu = 0.5$ s/m, $\alpha = 1.5$, $v_{\rm in} = 0.5$ m/s. At each corresponding point of the two parallel plots, the product $\mu v_{\rm in}$ is the same. It can be noted that, when not using any corrections, the weight of μ on the error is clearly greater than that of $v_{\rm in}$. (a) Maximum deviation on H as μ varies. (b) Maximum deviation on H as $v_{\rm in}$ varies.

that is the respective errors, as defined in (27) and (28), are equal to zero.[4]

As for the energy state upon the end of the interaction, the error on H_τ depends either on the error introduced by $\tilde{v}_{\rm out}$ (when the output velocity constraint described in Section IV-A2 is actually applied), or on the error introduced by the plain numerical simulation. As already stated in Section IV-A2, in the first case the maximum error on $v_{\rm out}$ is predictable, and clearly the same goes for the error on H_τ.

Equation (27) allows to quantitatively assess the improvements on the energy consistency of the numerical simulations.

The trend of error on H resulting from simulations with and without corrections is depicted in Fig. 11(a) and 11(b), which show that even the best numerical method among those considered (i.e., RK4) can be improved, especially for critical values of parameters (see Section III-B2) such as $\mu v_{\rm in} \to 0_+$.

2) Sequence of Impacts: In order to better appreciate the importance of the proposed corrections, a sequence of rebounds has been implemented applying a conservative force (e.g., gravity) during free motion only.[5] Thanks to this setup, one can track the accumulation of energy anomalies at each contact interaction.

To this end, the residual energy $H^{\rm sim}_{\tau(i)}$ of numerical simulations after the ith impact was examined and compared to the residual energy $H_{\tau(i)}$ due to the exit velocity $v_{\rm out}(i)$ of the ith rebound, computed numerically as a zero of (6). The corresponding relative error provides a measure of the inaccuracies *accumulated* during the sequence of impacts. In addition, the deviation of $H^{\rm sim}$ with respect to the analytical curve $H(v)$ was measured according to (27) along the whole sequence of impacts.

Table III shows the errors on the residual energy $H^{\rm sim}_{\tau(100)}$ and the maximum deviations of $H^{\rm sim}$ occurred during a sequence of 100 impacts, for simulation examples following *case 1* and *case 2*. Notice that, since in some simulations the energy can strongly oscillate during contact (see Fig. 7), the maximum deviations from $H(v)$ do not necessarily reflect the accumulated errors. The last columns show the errors resulting from simulations corrected as suggested in Section IV-A, where the error threshold for the output velocity constraint is set to 0.1%, i.e., the correction is always applied (see Fig. 3).

In order to better understand the importance of such corrections, Fig. 12 provides a partial plot of the sequence of rebounds for *case 2*, where only the two better performing simulations are depicted.

[4]Not taking into account the inherent errors related to the representation of numbers in the digital domain.

[5]That is, the external force is temporarily suspended during the short contact at each rebound. As shown in Section II-B, when an external force is applied during contact, no closed-form analytical result is available, this way making the corrections provided in Section IV-A unsuitable.

TABLE III
SUMMARY OF ERRORS ON THE RESIDUAL ENERGY $H^{\mathrm{sim}}_{\tau(100)}$ (ACCUMULATED ERROR), AND MAXIMUM DEVIATIONS OF H^{sim} CALCULATED ACCORDING TO (27) ALONG THE WHOLE SUCCESSION OF 100 REBOUNDS. (a) SIMULATION EXAMPLES FOLLOWING *case 1*. THE VALUES OF PARAMETERS ARE THE SAME AS IN FIG. 7, EXCEPT FOR INPUT VELOCITIES DECREASING AT EACH REBOUND. (b) SIMULATION EXAMPLES FOLLOWING *case 2*. THE VALUES OF PARAMETERS ARE THE SAME AS IN FIG. 8, EXCEPT FOR INPUT VELOCITIES DECREASING AT EACH REBOUND

(a)

errors on H	AM1	Verlet	Heun	RK4	corrected sim.
accum. %err	12.022	10.059	147.036	0.907	$< 10^{-7}$
max %dev	72.962	72.107	63.043	14.466	0

(b)

errors on H	AM1	Verlet	Heun	RK4	corrected sim.
accum. %err	24.285	15.780	946.600	2.153	0.001
max %dev	69.156	43.966	27.418	6.255	0

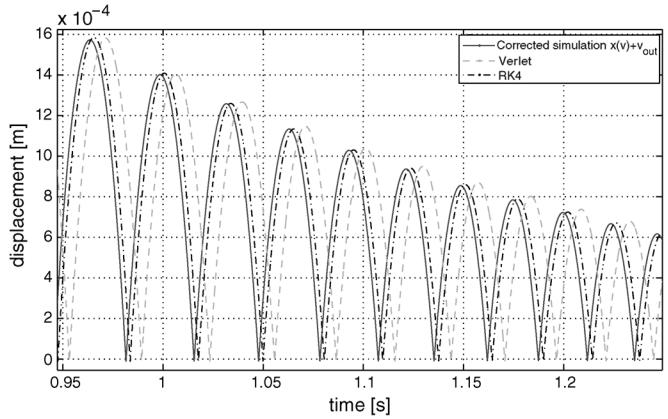


Fig. 12. Sequence of rebounds obtained from Verlet- and RK4-based simulations following *case 2*, compared with the trajectory of a corrected simulation.

From Table III and Fig. 12 it is evident that, even in case of errors apparently negligible for a single impact (see the errors relative to RK4 in Table II), indeed energy inconsistencies accumulate and can become noticeable—if not disastrous—after a certain number of contact interactions.

On the other hand, the errors resulting from the corrected simulations are actually negligible, and they are only due to inaccuracies in the forced exit velocity $\tilde{v}_{\mathrm{out}}$. The effectiveness of such corrections is confirmed comparing the results of a $16\times$ oversampled RK4 simulation used as a reference: in *case 1* the accumulated error on H is 0.001% and the maximum deviation is 0.024%, while in *case 2* they are, respectively, 0.002% and 0.006%.

C. Computational Cost

In this section, the computational costs of both the numerical implementations seen in Section III-A and the corrections suggested in Section IV-A are taken into account. In particular, the cost has been measured as the number of operations (i.e., memory write/read accesses and arithmetical operations) needed to execute an algorithm.

Table IV(a) summarizes the cost of the algorithms implementing the numerical methods. Since the AM1-based implementation makes use of Newton's method, its cost is displayed on two sub-columns: the first one shows the constant cost per sample, while the second one (in italics) shows the cost of a single iteration of Newton's method. Notice that the number of iterations per sample is not predictable.

TABLE IV
NUMBER OF OPERATIONS NEEDED BY THE NUMERICAL METHODS SHOWN IN SECTION III-A AND THE CORRECTIONS DESCRIBED IN SECTION IV-A. THE TOTALS IN BRACKETS ACCOUNT FOR WRITE/READ OPERATIONS. (a) SINCE THE AM1-BASED IMPLEMENTATION MAKES USE OF NEWTON'S METHOD, ITS COST IS DISPLAYED ON TWO SUB-COLUMNS: THE FIRST ONE SHOWS THE COST PER SAMPLE, WHILE THE SECOND ONE (IN ITALICS) SHOWS THE COST OF A SINGLE NEWTON'S METHOD ITERATION. (b) THE LAST COLUMN (IN ITALICS) SHOWS THE COST OF A SINGLE ZERO-FINDING ITERATION ON (6)

(a)

	AM1		Verlet	Heun	RK4
write	6	*10*	5	5	20
read	16	*33*	20	21	72
$+/-$	3	*8*	7	8	22
$\times$	8	*11*	5	5	18
$\div$	0	*1*	3	3	8
bit-shift	0	*0*	2	3	6
exp	0	*1*	1	1	4
log	0	*0*	0	0	0
compare	0	*4*	1	1	4
TOTAL	11 (33)	*25 (68)*	19 (44)	21 (47)	62 (154)

(b)

	hybrid correct.	$\tilde{v}_{\mathrm{out}}$ constraint	v_{out} as zero of (6)
write	1	2	*1*
read	7	4	*3*
$+/-$	6	5	*4*
$\times$	7	6	*3*
$\div$	3	4	*1*
bit-shift	0	1	*0*
exp	0	1	*0*
log	1	0	*1*
compare	1	1	*0*
TOTAL	18 (26)	18 (24)	*9 (12)*

Table IV(b) shows the cost of the corrections. In particular, the columns "hybrid correct." and "$\tilde{v}_{\mathrm{out}}$ constraint" account for straightforward implementations of (3) and (7), (8), respectively, which consist in substituting the continuous-time variables with their numerical counterparts.

For comparison, the last column reports the cost of a single zero-finding iteration on (6). The number of iterations depends on μ and v_{in}, and is in the order of some tens (usually between 15 and 40). Moreover, as μ decreases the number of iterations increases. It is clear that, despite being more precise than the approximate value $\tilde{v}_{\mathrm{out}}$, the value computed numerically as a zero of (6) implies several times the number of operations required by $\tilde{v}_{\mathrm{out}}$.

Since the computational load of simple write/read operations is generally low (if not negligible), two totals for each column are reported: one excluding write/read operations and, in brackets, one accounting for them.

Recalling that the hybrid correction only affects the computational cost during contact, whereas the output velocity constraint is applied at most once per impact event (as shown in Section IV-A2, the output velocity constraint should be applied conditionally), from Table IV one can infer that Verlet- or Heun-based simulations with corrections are roughly three times as efficient as plain RK4-based simulations during free motion, and almost twice as efficient during contact. On the other hand, while the exact computational load of an AM1-based simulation

is not predictable, it can be noted that it already matches the cost of a RK4-based one after two iterations of Newton's method.

D. Evaluation of Methods

The considered implementations can be finally evaluated in the light of the results regarding their accuracy (see Section III-B and IV-B2) and computational load (see Section IV-C).

It can be stated that, among the non-corrected implementations, the best all-round performance is achieved by RK4-discretized systems, although they are quite computationally expensive. On the other hand, a corrected Verlet-discretized system is generally at least as good as a non-corrected RK4 implementation, at a fraction of its computational load.

As for AM1-based implementations, it was shown that they behave quite poorly in the critical regions identified as *case 1* and *case 2*. These poor results in terms of accuracy, together with a generally high (and non-predictable) computational load, set the AM1 method as a hardly recommendable choice. Apparently such conclusion can be quite surprising, especially if one considers that AM1 is the only implicit and A-stable method among those considered. However, the presence of nonlinearities, together with the inaccuracies introduced by the K-method and Newton's method, justify the behavior of AM1-based implementations.

V. Conclusion

A nonlinear physical model of impact with sound synthesis applications has been reviewed, and its properties have been studied using both analytical tools and numerical simulations.

Several numerical realizations have been compared, and their shortcomings with regard to the corresponding analytical results have been pointed out. Special emphasis has been placed on energy consistency.

It has been shown that by exploiting the analytical results provided, the inconsistencies of the numerical realizations can be amended, thus restoring the correct energy state of the simulated systems, during and after contact.

Future research will consider finding a closed-form approximation of the release velocity for the system of (21), where a constant external force is applied, in this way allowing to implement suitable corrections. Even without such a closed-form solution, zero-finding numerical procedures could be profitably used to fix individual impacts. More interestingly for applications in acoustics, solutions will be investigated for extending the corrections to the case where vibrational losses are present, thus being applicable to impacts with resonating objects.

Appendix

A. Polynomial Expansion of the Output Velocity

As mentioned in Section II-A2, Hunt and Crossley [12] suggested that, in the limit $\mu v_{\rm in} \to 0_+$, the restitution coefficient can be approximated by the linear function $\widetilde{E}(\mu v_{\rm in}) = (1 - 2/3\mu v_{\rm in})$. Then, recalling (5), the corresponding output velocity is

$$\tilde{v}_{\rm out}(\mu, v_{\rm in}) = \left(-v_{\rm in} + \frac{2}{3}\mu v_{\rm in}^2\right). \tag{29}$$

This result can be easily verified through a Taylor expansion of the two sides in (6), using the approximation $\log|1+\epsilon| \approx \epsilon - (\epsilon^2/2) + (\epsilon^3/3)$, for $\epsilon \to 0_+$.

The same approach can be used in order to find a polynomial expansion of order n

$$\tilde{v}_{\rm out} = \sum_{j=1}^{n} a_j \cdot v_{\rm in}^j \tag{30}$$

in the limit $\mu v_{\rm in} \to 0_+$. Equation (30) is then substituted into the left-hand side of the Taylor expansion of (6). For the case $n = 4$ one obtains

$$\left[\frac{\mu^2}{2}v_{\rm out}^2 - \frac{\mu^3}{3}v_{\rm out}^3 + \frac{\mu^4}{4}v_{\rm out}^4 - \frac{\mu^5}{5}v_{\rm out}^5\right]_{v_{\rm out}=\sum_1^4 a_j v_{\rm in}^j} = \frac{1}{2}\mu^2 v_{\rm in}^2 - \frac{1}{3}\mu^3 v_{\rm in}^3 + \frac{1}{4}\mu^4 v_{\rm in}^4 - \frac{1}{5}\mu^5 v_{\rm in}^5. \tag{31}$$

The coefficients a_j are then determined by equating the two sides of (31) term by term, leading to the system

$$\begin{cases} \frac{1}{2}\mu^2 a_1^2 = \frac{1}{2}\mu^2 \\ \frac{1}{3}\mu^3 a_1^3 - \mu^2 a_1 a_2 = \frac{1}{3}\mu^2 \\ \frac{1}{4}\mu^4 a_1^4 - \mu^3 a_1^2 a_2 + \mu^2 a_1 a_3 + \frac{1}{2}\mu^2 a_2^2 = \frac{1}{4}\mu^2 \\ \frac{1}{5}\mu^5 a_1^5 - \mu^4 a_1^3 a_2 + \mu^3 a_1 a_2^2 + \mu^3 a_1^2 a_3 - \\ \quad \mu^2 a_1 a_4 - \mu^2 a_2 a_3 = \frac{1}{5}\mu^2 \end{cases} \tag{32}$$

which yields

$$a_1 = -1, \quad a_2 = \frac{2}{3}\mu, \quad a_3 = -\frac{4}{9}\mu^2, \quad a_4 = \frac{44}{135}\mu^3. \tag{33}$$

It follows that, in (30), $v_{\rm out}$ is a function of $(\mu, v_{\rm in})$ only. Equations (30) and (33) result in a slightly better approximation than (29), but still provide a local approximation.

B. Nonlocal Approximation for the Output Velocity

From (6) $v_{\rm out}$ can be written as

$$v_{\rm out} = \frac{1}{\mu}\left[(1+\mu v_{\rm in})e^{-\mu\Delta v} - 1\right] = v_{\rm lim}\left[1 - (1+\mu v_{\rm in})e^{-\mu\Delta v}\right] \tag{34}$$

where $\Delta v = v_{\rm in} - v_{\rm out}$. Equation (34) emphasizes that the convergence $v_{\rm out} \to v_{\rm lim}$ for $v_{\rm in} \to +\infty$ is governed by the fast-decreasing term $e^{-\mu\Delta v}$.

For small $v_{\rm in}$'s, the zeroth order approximation of the exponential term is $e^{-2\mu v_{\rm in}}$; however, it is easy to verify that the equation

$$\tilde{v}_{\rm out} = v_{\rm lim}\left[1 - (1+\mu v_{\rm in})e^{-2\mu v_{\rm in}}\right] \tag{35}$$

does not provide an accurate approximation of $v_{\rm out}$ for small $v_{\rm in}$'s. More precisely, its Taylor expansion around $v_{\rm in} = 0$ only matches the first Taylor coefficient of the expansion (30)

$$\tilde{v}_{\rm out} \underset{v_{\rm in}\to 0_+}{\approx} \sum_{j=1}^{n} \tilde{a}_j \cdot v_{\rm in}^j, \quad \tilde{a}_1 = -1 \quad \text{and} \quad \tilde{a}_j \neq a_j \quad (j > 1). \tag{36}$$

This qualitative discussion justifies to some extent the nonlocal approximation for $v_{\rm out}$ given in (7). Indeed (7) combines the two views—$\mu v_{\rm in} \to 0_+$ and $\mu v_{\rm in} \to +\infty$—on $v_{\rm out}$: the exponential term ensures the convergence $v_{\rm out} \to v_{\rm lim}$ for high values of

μv_{in}, while the polynomial coefficients b_j are determined by imposing that the Taylor expansion of (7) around $v_{\text{in}} = 0$ matches that of (30). Recalling that

$$e^{-2\mu v_{\text{in}}} \underset{v_{\text{in}} \to 0_+}{\approx} \sum_{j=1}^{n} c_j \cdot v_{\text{in}}^j, \quad c_j = \frac{(-1)^j}{j!}(2\mu)^j \tag{37}$$

then one can verify that the lth-order coefficient a_l of the Taylor expansion of (7) is given as $a_l = (1/\mu)\sum_{j=1}^{l} b_j \cdot c_{(l-j)}$. Then the coefficients b_j can be determined recursively as

$$b_0 = 1, \quad b_j = \mu a_j - \sum_{l=0}^{j-1} b_l \cdot c_{(j-l)} \quad (1 \le j \le n). \tag{38}$$

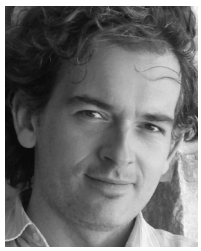
Applying this recursive equation for the case $n = 4$ yields (8).

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their insightful and helpful comments.

REFERENCES

[1] D. W. Marhefka and D. E. Orin, "A compliant contact model with nonlinear damping for simulation of robotic systems," *IEEE Trans. Syst., Man, Cybern. A: Syst. Humans*, vol. 29, no. 6, pp. 566–572, Nov. 1999.
[2] M. Mahvash, V. Hayward, and J. Lloyd, "Haptic rendering of tool contact," in *Proc. Eurohaptics*, 2002, pp. 110–115.
[3] M. Müller, J. Stam, D. James, and N. Thürey, "Real time physics: Class notes," in *Proc. SIGGRAPH '08: ACM SIGGRAPH 2008 Classes*, New York, 2008, pp. 1–90, ACM.
[4] A. Chaigne and J. Kergomard, *Acoustique des Instruments de Musique*. Paris, France: Belin, 2008.
[5] F. Avanzini and D. Rocchesso, "Modeling collision sounds: Non-linear contact force," in *Proc. Int. Conf. Digital Audio Effects (DAFx-01)*, Limerick, Ireland, Dec. 2001, pp. 61–66.
[6] M. Rath and D. Rocchesso, "Continuous sonic feedback from a rolling ball," *IEEE Multimedia*, vol. 12, no. 2, pp. 60–69, Apr./Jun. 2005.
[7] D. Rocchesso and F. Fontana, Eds., The Sounding Object Mondo Estremo, 2003 [Online]. Available: http://www.soundobject.org/
[8] P. Flores, J. P. Claro, and H. M. Lankarani, *Kinematics and Dynamics of Multibody Systems With Imperfect Joints: Models and Case Studies*. New York: Springer, 2008.
[9] G. Kuwabara and K. Kono, "Restitution coefficient in a collision between two spheres," *Jap. J. Appl. Phys.*, vol. 26, no. 8, pp. 1230–1233, 1987.
[10] A. Stulov, "Dynamic behavior and mechanical features of wool felt," *Acta Mechanica*, vol. 169, no. 1, pp. 13–21, 2004.
[11] L. Vu-Quoc and X. Zhang, "An elastoplastic contact force-displacement model in the normal direction: Displacement-driven version," in *Proc.: Math., Phys., Eng. Sci.*, 1999, pp. 4013–4044.
[12] K. H. Hunt and F. R. E. Crossley, "Coefficient of restitution interpreted as damping in vibroimpact," *ASME J. Appl. Mech.*, pp. 440–445, June 1975.
[13] N. Diolaiti, C. Melchiorri, and S. Stramigioli, "Contact impedance estimation for robotic systems," *IEEE Trans. Robotics*, vol. 21, no. 5, pp. 925–935, Oct. 2005.
[14] H. M. Lankarani and P. E. Nikravesh, "A contact force model with hysteresis damping for impact analysis of multibody systems," *J. Mech. Design*, vol. 112, no. 3, pp. 369–376, 1990.
[15] P. Flores, J. Ambrósio, J. Claro, and H. Lankarani, "Influence of the contact-impact force model on the dynamic response of multi-body systems," *Proc. Inst. Mech. Eng.-K*, vol. 220, no. 1, pp. 21–34, 2006.
[16] L. Pust and F. Peterka, "Impact oscillator with Hertz's model of contact," *Meccanica*, vol. 38, no. 1, pp. 99–116, 2003.
[17] S. Papetti, F. Fontana, M. Civolani, A. Berrezag, and V. Hayward, "Audio-tactile display of ground properties using interactive shoes," in *Haptic and Audio Interaction Design*, ser. Lecture Notes in Computer Science, R. Nordahl, S. Serafin, F. Fontana, and S. Brewster, Eds. Berlin/Heidelberg, Germany: Springer, 2010, vol. 6306, pp. 117–128 [Online]. Available: http://dx.doi.org/10.1007/978-3-642-15841-4_13
[18] A. Fettweis, "Wave digital filters: Theory and practice," *Proc. IEEE*, vol. 74, no. 2, pp. 270–327, Feb. 1986.
[19] S. Bilbao, "Robust physical modeling sound synthesis for nonlinear systems," *IEEE Signal Process. Mag.*, vol. 24, no. 2, pp. 32–41, Mar. 2007.
[20] A. Sarti and G. De Poli, "Toward nonlinear wave digital filters," *IEEE Trans. Signal Process.*, vol. 47, no. 6, pp. 1654–1668, Jun. 1999.
[21] I. Millington, *Game Physics Engine Development*. San Mateo, CA: Morgan Kaufmann, 2007.
[22] J. E. Colgate and J. M. Brown, "Factors affecting the Z-width of a haptic display," in *Proc. IEEE Int. Conf. Robot. Autom.*, San Diego, CA, May 1994, pp. 3205–3210.
[23] K. Kuchenbecker, J. Fiene, and G. Niemeyer, "Improving contact realism through event-based haptic feedback," *IEEE Trans. Vis. Comput. Graphics*, vol. 12, no. 2, pp. 219–230, Mar.–Apr. 2006.
[24] S. Bilbao, *Numerical Sound Synthesis*. New York: Wiley, 2009.
[25] F. Avanzini and D. Rocchesso, "Physical modeling of impacts: Theory and experiments on contact time and spectral centroid," in *Proc. Int. Conf. Sound Music Comput. (SMC'04)*, 2004, pp. 287–293.
[26] B. L. Giordano, D. Rocchesso, and S. McAdams, "Integration of acoustical information in the perception of impacted sound sources: The role of information accuracy and exploitability," *J. Exper. Psychol.: Human Percept. Perf.*, vol. 36, no. 2, pp. 462–476, 2010.
[27] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics*, 2nd ed. New York: Springer, 2007.
[28] G. Borin, G. De Poli, and D. Rocchesso, "Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 597–605, Sep. 2000.

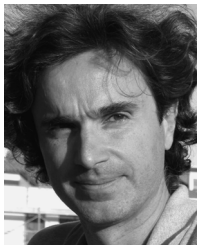


Stefano Papetti received the Laurea degree in computer engineering from the University of Padova, Padua, Italy, in 2006, and the Ph.D. degree in computer science from the University of Verona, Verona, Italy in 2010.

During the Ph.D. degree, he worked as a Visiting Researcher at Medialogy (Department of Architecture, Design and Media Technology), Aalborg University, Copenhagen, Denmark. Since 2006, he has been with the Computer Science Department, University of Verona, where he currently works as Research Associate. His research activity is focused on models and applications for interactive sound synthesis, and on the design and evaluation of auditory and haptic interfaces.

Federico Avanzini received the Laurea degree (*cum laude*) in physics from the University of Milano, Milan, Italy, in 1997 and the Ph.D. degree in information engineering from the University of Padova, Padua, Italy, in 2001 with a research project on sound and voice synthesis by physical modeling.

During the Ph.D. degree, he also worked as a Visiting Researcher at the Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, Espoo, Finland. Since 2002, he has been with the Sound and Music Computing Group, Department of Information Engineering, University of Padova, where he is currently an Assistant Professor, teaching courses in computer science and sound and music computing. His main research interests are in the area of sound synthesis and processing, with particular focus on musical sound synthesis, nonspeech sound in multimodal interfaces, voice synthesis, and analysis. He has authored more than 70 publications in peer-reviewed international journals and conferences, and has been involved in numerous EU research projects. He is PI of the EU project DREAM (Culture2007) and of industry-funded projects.



Davide Rocchesso received the Ph.D. degree from the University of Padova, Padua, Italy, in 1996.

From 1998 and 2006, he was with the Computer Science Department, University of Verona, Verona, Italy, as an Assistant and Associate Professor. Since 2006, he has been with the Department of Art and Industrial Design, IUAV University of Venice, Venice, Italy, as an Associate Professor. He has been the coordinator of EU project SOb (the Sounding Object), and local coordinator of the EU project CLOSED (Closing the Loop Of Sound Evaluation and Design) and of the Coordination Action $\mathrm{S2S}^2$ (Sound-to-Sense; Sense-to-Sound). He authored or coauthored over one hundred publications in scientific journals, books, and conferences. His main research interests are sound modelling for interaction design, sound synthesis by physical modeling, and design and evaluation of interactions.

Dr. Rocchesso is currently chairing the COST Action IC-0601 SID (Sonic Interaction Design).