Antialiasing for Nonlinearities: Acoustic Modeling and Synthesis Applications

Harvey Thornburg

Center for Computer Research in Music and Acoustics

Stanford University

harv23@ccrma.stanford.edu

ABSTRACT

Nonlinear elements have manifold uses in acoustic modeling, audio synthesis and effects design. Of particular importance is their capacity to control oscillation dynamics in feedback models, and their ability to provide digital systems with a natural overdrive response. Unfortunately, nonlinearities are a major source of aliasing in a digital system. In this paper, alias suppression techniques are introduced which are particularly tailored to preserve response dynamics in acoustic models. To this end, a multirate framework for alias suppression is developed along with the concept of an aliasing signal-to-noise ratio (ASNR). Analysis of this framework proceeds as follows: first, relations are established between ASNR vs. computational cost/delay given an estimate of the reconstructed output spectrum; second, techniques are given to estimate this spectrum in the worst case given only a few statistics of the input (amplitude, and bandwidth). These tools are used to show that hard circuit elements (*i.e.* saturator, rectifier, and other piecewise linear systems found in bowed-string and single-reed instrument models) generate significant ASNR given reasonable computational constraints. To solve this problem, a parameterizable, general-purpose method for constructing monotonic "variably soft" approximations is developed and demonstrated to greatly suppress aliasing without additional computational expense. The monotonicity requirement is sufficient to preserve response dynamics in a variety of practical cases.

1. INTRODUCTION

A worthwile goal in music synthesis is to design instruments that convey expression in performance. To this end, an important aspect is that the *energy* of performance be reflected in the sound. By overblowing a flute, for example, the performer can navigate complex regimes of oscillation, as well as effect subtle tonal variations within the regimes.

A linear system is insufficient for this task; moreover, if the instrument is used as a source, it must be capable of sustained oscillation, implying feedback. In a landmark paper[1], McIntyre, Schumacher and Woodhouse demonstrate that virtually all oscillatory systems found in musical acoustics generalize to a nonlinear feedback model in which a passive, linear filter is connected in feedback around an active nonlinear part. In the scope of this paper, the nonlinearity is considered memoryless. The nonlinearity determines primary characteristics (which instrument) whereas the linear part (delays and lowpass filter) influences secondary characteristics such as pitch and timbral nuance.

Steady-state timbre is often nonsinusoidal, suggesting two kinds of dynamic balance involving linear and nonlinear parts. First is the classical "energy balance" responsible for limit cycle (oscillatory) behavior. The linear part dissipates energy supplied by the performer through the nonlinearity. The level of oscillations is governed by this equilibrium * Of additional importance is the spectral information balance. Lowpass elements in the linear part restrict bandwidth whereas the nonlinear part tends to expand bandwidth. If the result is to be oscillatory rather than chaotic, the evolution of timbre is seen to depend greatly on the nature of this balance.

In short, the nonlinear feedback characterization is vital to the phenomenon of expressive control. However, the nonlinear part produces aliasing in digital implementation. This is especially true of the hard "circuit nonlinearities" (*e.g.* saturator, rectifier, dead-zone, pulse, etc.) as they produce discontinuities and corners in the output. Aliasing is especially problematic in closed-loop systems, as it may propagate around the loop or otherwise disrupt the spectral information balance, suppressing regimes of oscillation or even displacing them towards chaos.

To reduce aliasing one can change the sampling rate (within the efficient multirate framework; see Section 2) as well as approximate the nonlinearity with an equivalent less conducive to aliasing. Approximations should be made, however, with the express objective of preserving as much of the *closed-loop* system dynamics as possible. This contrasts with polynomial approximation techniques, which favor the output spectrum (an *open-loop* characteristic). Ideally one aims to duplicate the exact locations of oscillatory and chaotic regimes and the quality of transitions between regimes. Unfortunately, direct closed-loop analyses prove forbiddingly complex. An efficient compromise is to ensure various *control properties* are maintained for the nonlinear part. This compromise is further justified in that hard circuit nonlinearities are often gross idealizations of the physical law, yet certain properties are maintained; among these are *small-signal equivalence*, *monotonicity*, and *preservation of asymptotes*.

*Alternatively, one can lump the small-signal linearization of the nonlinear part as a gain factor in the linear part. Here the linear part is unstable, resulting in whereas the nonlinear part reflects limitations of the energy supply. This perspective is common, for instance, when modeling circuits such as the self-oscillating voltage-controlled filter, or other situations where the role of the performer *viz.* energy transfer is indirect.

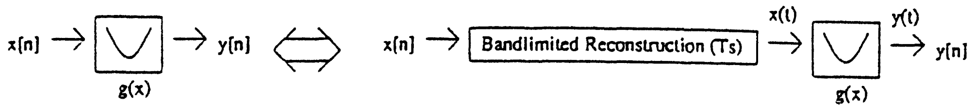

Figure 1. System equivalents for memoryless digital nonlinearity $g(x)$.

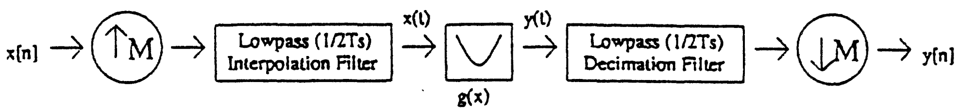

Figure 2. Multirate framework for aliasing suppression

2. A MULTIRATE FRAMEWORK FOR ALIAS SUPPRESSION

To motivate the multirate framework, it is necessary to consider how digital nonlinearities produce aliasing. Let $g(x)$ be a memoryless nonlinear map. Figure 1 gives system equivalents. In the figure $y(t)$ is generally not bandlimited to Nyquist $(T_s/2)$. What is missing is the antialiasing filter between $y(t)$ and the sampler. To implement this filter requires an infinite sampling rate, as all frequencies exceeding Nyquist must be rejected to prevent them from folding over. This is not available so successive approximations using a multirate framework (Fig. 2) are considered.

Portions of the output spectrum rejected by the decimation filter are shown in Fig. 3. As upsampling factor M increases, regions '2' expand; regions '3' shift along the spectrum but do not change size.

By using polyphase structures [3], computational cost of the multirate framework is kept linear in M. As closed-loop response is extremely sensitive to phase, interpolation/decimation filters should be linear phase, which adds a (possibly fractional) number of unit delays to the loop at base sampling rate T_s. In waveguide models, it is easier to adjust for extra delays (affecting pitch) than it is for an arbitrary nonlinear phase response. Delay compensation for lumped models, if possible, proceeds by root-locus method applied to the linearized equivalent; see [2] for details.

To quantify aliasing performance gains, it is natural to define an *aliasing signal-to-noise ratio*, treating foldover as noise and the unaliased portion as signal:

$$ASNR(M,x) \triangleq \frac{\int_0^{1/(2T_s)} |Y(f)|^2 df}{\sum_{k=1}^{\infty} \int_{(kM-1/2)/T_s}^{(kM+1/2)/T_s} |Y(f)|^2 df} \qquad (1)$$

Here dependence on the input ($x[n]$) is made explicit. The aliasing *cost* (denoted $J_{alias}(M,x)$) is taken as the reciprocal of $ASNR(M,x)$. By varying M one can monitor a cost vs. computations curve. The aim in approximation is to design $g(x)$ such that the tradeoff ($J_{alias}(M,x)$ vs. M) is Pareto optimal. The optimality is *viz.* constraints on $g(x)$; namely, the control properties discussed in Section 1. Furthermore, one requires that $g(x)$ be somehow "close" to $g_0(x)$, the latter representing the ideal nonlinearity. It is useful to produce a *family* of approximations, subject to control property constraints, in which Pareto optimality over all criteria (aliasing cost, computational cost, closeness to g_0) is approached.

The above specification is still incomplete, as $J_{alias}(M,x)$ depends on $x[n]$. It is natural to minimize aliasing cost for the worst case input, as to minimize restrictions on other parts of the system. The cost may be unbounded in lieu of additional constraints on the input. Two natural constraints are amplitude and bandwidth: amplitude is fixed by the system's dynamic range; bandwidth, if unknown, is fixed at Nyquist. The worst-case input depends on the nonlinearity; it can be obtained as the solution to a separate optimization problem. In practice, a useful proxy is found to be a cosine at maximum amplitude with frequency set to the bandwidth limit.

Formulas for the output spectrum $Y(f)$ are now given, with and without the simplifying assumption of a cosine input. The formulas fall into two classes, one based on Taylor series approximation of the nonlinearity; the other based on Fourier analysis of the output and/or decomposition of the nonlinear map into phase modulation (PM) operators. The Taylor series approach is limited for hard nonlinearities due to the analyticity requirement; however insights may be gained:

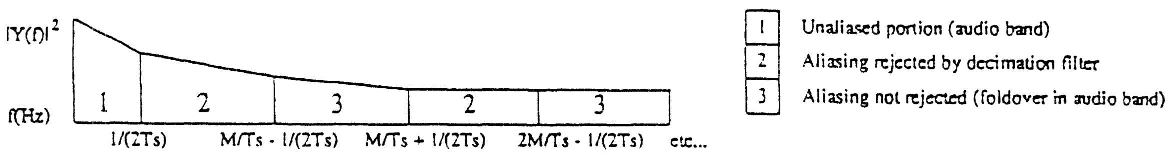

Figure 3. Output spectrum: aliasing rejection and foldover regions

$$Y(f) = \sum_{k=0}^{\infty} \frac{1}{k!} g^{('k)}(0)\, [X(f)]^{(*k)} \tag{2}$$

Here $('k)$ denotes repeated differentiation and $(*k)$ iterated convolution. For a cosine input $x(t) = A\cos(\omega_0 t)$, the iterated convolution obtains a closed form solution via binomial expansion:

$$[X(f)]^{(*k)} = A^k \sum_{l=0}^{k} \binom{k}{l} \delta(2\pi(f-(k-2l)f_0)) \tag{3}$$

Equation (3) is recognized as a special case of the bandwidth expansion formulas given by Steer [4] for Volterra kernels with sinusoidal input. It should also be remarked that if $g(x)$ is a finite order polynomial, the sum in (2) will be finite. The output bandwidth is $(k-1)/2T_s$; choosing $M > (k+1)/2$ eliminates aliasing. Difficulty lies in that hard circuit nonlinearities are neither analytic nor are analytic approximants useful when restricted to finite order. Furthermore, an implicit definition $g_y(y) = g_x(x)$ may prove advantageous, for which determining the Taylor expansion is analytically and computationally prohibitive.

For a sinusoidal input, the output of a memoryless nonlinearity is periodic. Thus, great simplification abounds when considering Fourier-based approaches, as there are few restrictions (*e.g*, Dirichlet and Riemann-integrability conditions) on the map $g(x)$. The Fourier spectral representation and projections for coefficients a_k are given:

$$Y(f) = \sum_{k=-\infty}^{\infty} a_k \delta(2\pi k(f-f_0)); \quad a_k = f_0 \int_0^{1/f_0} g(A\cos(2\pi f_0 t)) e^{-2\pi i k f_0 t} dt \tag{4}$$

It is arguably more useful to obtain the output spectrum in terms of the map $g(x)$ itself, or its Fourier transform $G(\omega)$. The latter approach follows the decomposition of g(x) into a spectrum of transcendental PM operators $e^{iA\cos\omega_0 t}$; the output Fourier series of a single operator being:

$$e^{iA\cos\omega_0 t} = \sum_{k=-\infty}^{\infty} i^k J_k(A) e^{ik\omega_0 t} \tag{5}$$

where $J_k(\cdot)$ denotes the Bessel function of integer order k.

Following [5], a straightforward computation involving (5) and the inverse transform relation $g(x) = \int_{-\infty}^{\infty} G(f) e^{i\omega x} d\omega$ obtains

$$a_k = \int_{-\infty}^{\infty} G(f) J_k(Af) df \tag{6}$$

Parseval's relation brings this into the amplitude domain:

$$a_k = \frac{1}{|A|} \int_{-A}^{A} g(x) j_k(\frac{x}{A}) dx; \quad j_k(x) \triangleq T_k(x)/\sqrt{1-x^2} \tag{7}$$

where $T_k(x) \triangleq \cos^{-1}(k\cos(x))$ is the k^{th} Chebyshev polynomial. The inverse Fourier transforms of the Bessel functions $j_k(x)$ are recognized as the dual basis for projections onto the space of Chebyshev polynomials on $(-1, 1)$. A more complete derivation is given in [6].

The uses of (7) are that a_k are given as inner products with $g(x)$ itself; numerical integration proceeds over a finite range, and intuition concerning the sensitivity of the output spectrum to the input *amplitude* is obtained. From ((4), (6), or (7)) and (1) one estimates the worst-case aliasing cost for any configuration of the multirate framework.

3. APPROXIMATION FRAMEWORK

Here practical solutions are given and results detailed concerning approximations within the multirate framework: Given a nonlinearity $g_0(x)$, based on physical measurements or idealizations thereof, it is desired to approximate with $g(x)$ such that aliasing is suppressed, subject to control property constraints. Goals are the threefold minimization of the costs: {a. $\|g(x) - g_0(x)\|$, where $\|\cdot\|$ is any norm; b. $\sup_{x(t)}\{J_{alias}(M, x)\}$; c. M.}. Commonly relevant control properties are: {a. *Small-signal gain equivalence*: $g'(0) = A_0$; b. *Monotonicity*: $g'(x) > 0$ for $x \in S_0$; $g'(x) \leq 0$ for $x \in S_1$; c. *Asymptotes*: $\lim_{x\to\infty} g(x) = L_0$; $\lim_{x\to-\infty} g(x) = L_1$; d. *Odd/even property*: $g(x) = g(-x)$, all x; etc....}

Although many objectives and constraints are linear (advantage of (7)), the aliasing cost objective is not convex: direct optimization is unwieldy without further restriction. Fortunately, a simple construction is available for common cases of hard circuit nonlinearities; it satisfies relevant control properties by design, and is shown to effect substantial improvements in aliasing suppression. The idea is to design a variably soft nonlinearity; "softness" being a single parameter navigating the tradeoff between closeness to the ideal map and reduction of aliasing cost at fixed M.

The variably soft construction is introduced by example of the saturator from which others are simply obtained. The inverse map $x = S_p^{-1}(y)$ is obtained, $p \in [1, \infty]$ being an integer softness parameter. Let $\gamma_p(y)$ be an odd function with

Saturator	$y = S_p(x)$	$x = y + y^{2p}\gamma_p(y)$
Dead Zone	$y = D_p(x)$	$y = x - S_p(x)$
Pulse	$y = P_p(x)$	$y = \frac{d}{dx}S_p(x)$
Saturating Rectifier	$y = R_p(x)$	$y = 1 + \frac{(S_p(x-1/2)-S_p(x+1/2))}{2}$

Table 1. Construction for variably soft nonlinear maps

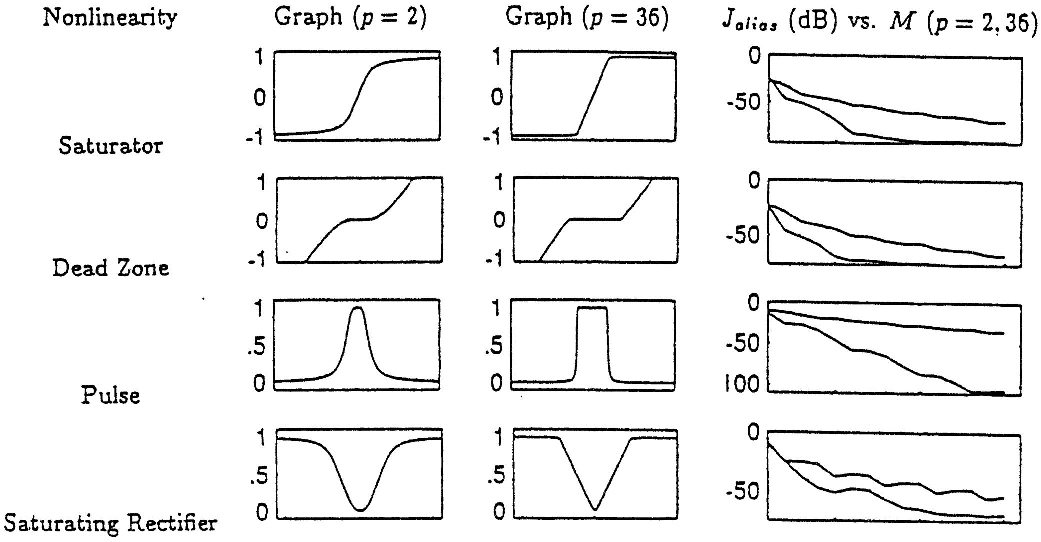

Figure 4. Aliasing Cost vs. Computations Tradeoff Comparisons

asymptotes $\gamma_p(\pm 1) = (\pm\infty)$. An arbitrary number of derivatives are zeroed (decreasing softness) without disturbing the odd-property or the asymptotes, by setting $x = y^{2p}\gamma_p(y)$. Finally the x-axis is sheared obtaining the correct inverse map: $x = S_p^{-1}(y) \triangleq y + y^{2p}\gamma_p(y)$. The forward map is computed in a table by iterating Newton's method.

Table 1 gives formulas for obtaining variably soft approximations for some circuit nonlinearities. Results and aliasing cost/computation curves are detailed in Fig. 4 for a sinusoidal input at frequency $1/8T_s$, amplitude 2.0, varying M from 1 to 16, using equations (1),(7) and $\gamma_p(y) = y/(1-y^{2p})$.

It is apparent from Fig. 4 that the variably soft construction greatly reduces aliasing while maintaining exactly the relevant control properties. For the saturator, when $M = 5$ there is a 22 dB gain in ASNR from $p = 36$ to $p = 2$; for the pulse (commonly used to model the bow-string scattering junction; after [7]) the difference is 27 dB, increasing sharply to 70 dB when $M = 14$.

4. CONCLUSIONS

A multirate framework for aliasing suppression is introduced along with a simple, "variably soft" construction for approximating ideal hard circuit nonlinearities such as to efficiently navigate the tradeoff between closeness to ideality and aliasing cost. The variably soft construction explicitly incorporates control property constraints (*e.g.* small-signal equivalence, monotonicity, asymptotes) which aim to preserve closed-loop response characteristics as opposed to merely the open-loop (spectral) properties of the approximation.

REFERENCES

1. M.E.Mcintyre, R.T.Schumacher, and J. Woodhouse, "On the oscillations of musical instruments," *J. Acoust. Soc. Am* 74 (5), pp. 1325–1345, 1983.
2. P.P.Vaidyanathan, *Multirate Systems and Filterbanks*, Prentice Hall, Englewood Cliffs, N.J., 1993.
3. T. Stilson and J.O.Smith, "Analysis of the moog vcf with considerations for digital implementation," *ICMC-96* , (Hong Kong), 1996.
4. M.B.Steer, P.J.Khan, and R.S.Tucker, "Relationship of volterra series and generalised power series," *Proc. IEEE* , pp. 1452–1454, Dec. 1983.
5. R. Gray and J. Goodman, *Fourier Transforms*, Kluwer Academic, Dordrecht, 1995.
6. H. Thornburg, "Derivation: Harmonic analysis via pm-decomposition," *unpublished manuscript* , 1998.
7. J.O.Smith, "Discrete-time modeling of acoustic systems with applications to sound synthesis of musical instruments," *Proc. Nord. Acoust. Meeting* , (Helsinki), 1996.