



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Implementação de Filtros Digitais para Afinação de Guias Digitais de Ondas

José Honório Rabelo Neto

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Orientador
Prof. Dr. Márcio da Costa Pereira Brandão

Brasília
2014

Universidade de Brasília — UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Bacharelado em Ciência da Computação

Coordenador: Prof. Dr. Homero Luiz Piccolo

Banca examinadora composta por:

Prof. Dr. Márcio da Costa Pereira Brandão (Orientador) — CIC/UnB
Prof. Dr. Alexandre Zaghetto — CIC/UnB
Prof. Dr. Camilo Chang Dórea — CIC/UnB

CIP — Catalogação Internacional na Publicação

Neto, José Honório Rabelo.

Implementação de Filtros Digitais para Afinação de Guias Digitais de Ondas / José Honório Rabelo Neto. Brasília : UnB, 2014.

41 p. : il. ; 29,5 cm.

Monografia (Graduação) — Universidade de Brasília, Brasília, 2014.

1. síntese, 2. áudio, 3. computação musical, 4. Interpolador de Lagrange, 5. guias digitais de ondas, 6. Karplus-Strong, 7. corda ideal

CDU 004.4

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro — Asa Norte
CEP 70910-900
Brasília-DF — Brasil

Agradecimentos

Agradeço primeiramente a Deus pela oportunidade de desfrutar em vida, e com saúde, das felicidades proporcionadas pela música. Agradeço à minha família pelo apoio incondicional, presentes a todo momento para reforçar e reafirmar meu comprometimento com este curso. Agradeço imensamente ao meu orientador, Prof. Dr. Márcio Brandão, que teve uma paciência sobrehumana por me acompanhar por esta tão longa jornada para a elaboração deste trabalho. E agradeço, com um carinho muito especial, à Fernanda, pelo companheirismo, paciência, superação e amor, imprescindíveis nestes últimos longos e estressantes meses de trabalho.

Resumo

Este trabalho trata da análise do processo de propagação de ondas transversais numa corda ideal e sua sintetização, através de técnicas de modelamento físico de instrumentos musicais, com o objetivo de produzir sons musicais afinados a uma altura arbitrária. Métodos simples de determinação de altura e afinação são propostos. São apresentadas as técnicas de Guias Digitais de Ondas e o algoritmo Karplus-Strong para a síntese de sons complexos de altura bem definida, com a utilização de filtros Interpoladores de Lagrange como soluções para sua afinação. A análise dos sons produzidos pelo protótipo implementado é apresentada ao final do trabalho.

Palavras-chave: síntese, áudio, computação musical, Interpolador de Lagrange, guias digitais de ondas, Karplus-Strong, corda ideal

Abstract

This paper concerns the analysis and synthesis of transversal wave propagation on an ideal string, using techniques for physical modeling of musical instruments, with the objective of producing sounds well tuned to an arbitrary pitch. Simple methods for pitch determination and tuning are proposed. Digital Waveguide and Karplus-Strong algorithm are presented as techniques for pitched sound synthesis, and Lagrange Interpolators as tuning solutions. The analysis of sounds produced by the implemented prototype is presented at the end of the paper.

Keywords: synthesis, audio, computer music, Lagrange Interpolator, Digital Waveguides, Karplus-Strong, ideal string

Sumário

1	Introdução	1
2	Sons Musicais	3
3	Guias Digitais de Ondas	6
3.1	Introdução	6
3.2	A Corda Vibrante	6
3.2.1	Equação da Corda Ideal	7
3.3	Realizações	11
3.3.1	Guias Digitais de Ondas	11
3.3.2	Algoritmo Karplus-Strong	12
4	Calibrando a Linha de Retardo	15
4.1	Introdução	15
4.2	Filtros Digitais	16
4.2.1	Interpolador de Lagrange	17
5	O Protótipo	20
5.1	Considerações iniciais	20
5.1.1	Modelo	21
5.2	Implementação	22
5.2.1	O excitador	22
5.2.2	Linha de retardo	23
5.2.3	Afinação	23
5.2.4	Usabilidade	23
5.3	Resultados	24
5.3.1	Metodologia	24
5.3.2	Análise	26
6	Conclusão	30
	Referências	31

Lista de Figuras

2.1	Frequência fundamental nas teclas do piano.	5
3.1	A corda vibrante ideal.	7
3.2	Forças atuantes num segmento da corda vibrante.	8
3.3	Modos de vibração de uma corda ideal.	9
3.4	Visualização do teorema de d'Alembert.	10
3.5	Topologia do algoritmo de Guias de Ondas Digitais.	11
3.6	Guia de Ondas Digitais com linhas de retardo consolidadas e terminações reflexivas.	12
3.7	Esquema do algoritmo Karplus-Strong.	13
3.8	Espectro de um som com frequência fundamental de $440Hz$, produzido pela técnica de Guias Digitais de Ondas.	14
3.9	Espectro de um som com frequência fundamental de $440Hz$, produzido pelo algoritmo Karplus-Strong.	14
4.1	Frequências obtidas por uma linha de retardo de comprimento inteiro. Considerou-se uma frequência de amostragem de $48000Hz$	15
4.2	Realização de um filtro FIR de ordem N em sua Forma Direta.	16
4.3	Função sinc atrasada em $D = 3$ e $D = 3.3$ amostras.	17
4.4	Interpolador de Lagrange de primeira ordem.	18
4.5	Interpolador de Lagrange de segunda ordem.	19
4.6	Interpolador de Lagrange de terceira ordem.	19
5.1	Diagrama-modelo implementado no protótipo.	21
5.2	Esquema do Excitador. Cada amostra é oferecida diretamente à saída.	22
5.3	Precisão unitária da linha de retardo e binária da FFT.	25
5.4	Detalhe da Figura 5.3.	26

Lista de Tabelas

5.1	Picos de amplitude para sons de alturas presentes no piano e extrapolação da última altura.	28
5.2	Aproximações de retardos de 0.5 amostras, correspondentes ao erro máximo de aproximação.	29

Capítulo 1

Introdução

Desde os tempos antigos, o homem se vê encantado com o fascínio misterioso que a música nos causa. Seja por simples contemplação da natureza ou para a sua própria sobrevivência, as informações percebidas por nosso sentido auditivo sempre foram de enorme importância. A partir do momento em que conseguimos certa estabilidade, e conseqüente alívio na luta pela sobrevivência, a música foi a conseqüência natural para a nossa espécie. O trabalho dos *luthiers* ganhou importância crescente ao longo da história, inicialmente para garantir maior uniformidade na produção de sons musicais e, posteriormente, para a própria evolução do conceito de produção musical.

A revolução cultural e científica possibilitada pela invenção da energia elétrica no século XIX também abraçou a música. No século seguinte, diferentes técnicas de síntese de áudio culminaram em instrumentos totalmente novos, como o MiniMoog, e o próprio computador. [1–3] Com a facilidade de implementação da Transformada Rápida de Fourier (*FFT*) e de filtros digitais, a relevância da computação na música cresceu enormemente através do desenvolvimento dos Processadores de Sinais Digitais (*DSPs*). O resultado foi uma avalanche de novas técnicas de síntese de áudio, das quais daremos destaque ao modelamento físico de instrumentos musicais. [4–15]

O modelamento físico consiste na descrição e simulação de características análogas à de instrumentos reais, como a tensão e densidade de uma corda, no ambiente computacional. São criados modelos matemáticos que aproximam os fenômenos físicos inerentes da propagação de ondas sonoras. Assim como um *luthier* procura a melhor combinação de materiais e formatos para obter um instrumento único, os modelos físicos miram na mesma direção através das equações que os compõem.

Este trabalho

Este trabalho trata da análise do processo de propagação de ondas numa corda e sua sintetização, através de técnicas de modelamento físico de instrumentos musicais. O objetivo é obter sons musicais afinados a uma altura arbitrária. O modelo de síntese consiste num misto entre as técnicas de Guias Digitais de Ondas e do algoritmo Karplus-Strong. Não se trata, portanto, de uma nova técnica de síntese, mas sim da análise e, principalmente, experimentação de técnicas amplamente utilizadas e estudadas no meio acadêmico para a produção de sons musicais.

O Capítulo 2 tem o objetivo de possibilitar a leitura do restante do trabalho por pessoas não-musicalizadas. Lá serão expostos, de forma breve, alguns dos conceitos e

definições a respeito de teoria musical, como intervalos, escalas, sons simples e complexos, a relação entre frequência e altura, dentre outros.

No Capítulo 3 o problema da corda vibrante será analisado. A equação de d'Alembert foi utilizada como forma de resolução da equação de onda, dada a sua facilidade de implementação em sistemas computacionais. Os algoritmos Karplus-Strong [5] e Guias Digitais de Ondas (*Digital Waveguides* [6]) serão analisados e discutidos, e servirão como base para o modelamento físico da corda.

Apesar de bastante simples e versáteis, por si só os algoritmos analisados são capazes de produzir apenas sons com alturas específicas. A limitação de produção de sons com período discreto, que constitui o problema a ser atacado pelo presente trabalho, será exposta e discutida no Capítulo 4, onde será proposta uma maneira de minimizá-la. Conforme citado na bibliografia da área, o emprego de filtros digitais para a aproximação de delays fracionários representa numa forma eficaz de contornar o problema, permitindo a afinação a uma escala musical arbitrária dos sons musicais produzidos [9, 16]. Em especial, filtros que implementam a Interpolação de Lagrange serão analisados.

Tudo isso culmina no Capítulo 5, onde será apresentado o protótipo implementado, bem como a apresentação e análise dos sons produzidos.

Capítulo 2

Sons Musicais

Numa proposta de síntese de áudio voltada para a produção de sons musicais, como é a presente, faz-se necessário o estabelecimento de regras que definam quais as condições em que determinada resposta em frequência é desejada, e qual não é. Neste capítulo alguns dos conceitos-chave são discutidos de forma breve, de maneira a possibilitar a leitura do restante do trabalho. Leituras posteriores são encorajadas para um entendimento mais profundo [17–19].

Sons simples e complexos

No contexto do presente trabalho, um *som simples* é aquele que pode ser descrito como uma função senoidal de apenas um fator. Em outras palavras, um som simples possui um espectro de frequência que assume um valor diferente de zero em um único ponto. Em contrapartida, um *som complexo* possui um espectro que apresenta mais de uma componente de frequência, constituindo uma combinação de dois ou mais sons simples.

Série harmônica

Consideremos o exemplo de uma corda vibrante com extremidades fixas. De forma natural, ela produz um som complexo com componentes de frequência da forma $N \cdot f$, sendo f um valor constante e N um número natural não-nulo. Tal característica também é prontamente observada em outros sistemas vibrantes, como em tubos acústicos. Cada uma destas componentes é definida como um *harmônico*, ou *sobretom* deste som complexo. A combinação destas componentes para a formação do som complexo foi um conceito herdado pela matemática para definição homônima.

Altura

A *altura* constitui-se uma característica do som intimamente relacionada à sua resposta em frequência, ainda que nem todos os sons possuam altura perceptivelmente definida [18]. Definimos grosseiramente os sons sem altura perceptível como *barulhos*, os quais não fazem parte do escopo deste trabalho.

No caso particular de sons simples, a sua resposta singular em frequência elimina ambiguidades na definição de sua altura, de modo que elas constituem um relacionamento funcional bijetor entre frequência e altura. Generalizando, a altura é percebida em sons

complexos cujo espectro de frequência revela uma predominância de componentes harmônicos (*sons harmônicos*) sobre outros componentes de frequência.

Com efeito, sons complexos com alturas bem definidas apresentam correspondência com a altura de sons senoidais simples e únicos [19]. No caso de sons complexos harmônicos como aqueles produzidos pelos modos de vibração de uma corda de extremidades fixas, a correspondência ocorre na região de frequência do primeiro harmônico. Desta forma, esta componente de frequência é chamada de *frequência fundamental*.

Timbre

Sons complexos de mesma altura podem apresentar diferentes sonoridades de acordo com as magnitudes de suas componentes de frequência. Em geral, cada instrumento possui a sua “assinatura” de resposta em frequência, à qual dá-se o nome de *timbre*. Esta característica marcante possibilita que um ouvinte identifique, por exemplo, que dois instrumentos diferentes estão produzindo a mesma nota simultaneamente.

Intervalos e Escalas

Apesar de fundamental em todo o processo de produção de sons musicais e da música propriamente dita, o conceito de altura, por si só, não é suficiente. Alguns sons, quando reproduzidos simultaneamente ou em sequência, são mais “agradáveis” de se ouvir que outros, constatado através de análises psicoacústicas [18]. Ao conjunto de relações de alturas que possuem maior afinidade entre si dá-se o nome de *escala musical*, a qual engloba *intervalos musicais* significativos. Um intervalo musical no contexto atual é definido pela razão entre as frequências fundamentais de dois sons complexos.

Em linhas gerais, intervalos harmônicos são comuns em várias dos exemplos de definições de escalas musicais. Os intervalos 1:1, 2:1 e 3:2, correspondentes aos três primeiros harmônicos de um som harmônico, são chamados na música ocidental de *uníssono*, *oitava* e *quinta*. Eles estão presentes na grande maioria das escalas documentadas e utilizadas na produção musical. Um destaque especial deve ser dado ao intervalo de oitava: notas musicais com alturas equivalentes às de frequências da forma $\lambda \cdot 2^N$, N natural e λ positivo constante, recebem o mesmo nome, tamanha a similaridade percebida pelo ouvido humano.

Discussão

Tomemos o piano padrão de 88 teclas como exemplo de instrumento musical. Usualmente, sua extensão vai de $27,5Hz$ até $4186,01Hz$, abrangendo pouco mais de 7 oitavas. A afinação usada é chamada de *temperamento igual*, na qual o intervalo musical entre duas notas adjacentes quaisquer é o mesmo para toda a extensão do instrumento. Cada oitava é dividida em 12 intervalos, chamados de *semitons*. Logo, a frequência fundamental f_N de cada tecla N é definida como

$$f_N = f_k \cdot 2^{\frac{N-k}{12}}, \{N, k\} \subset \mathbb{N}^*, \quad (2.1)$$

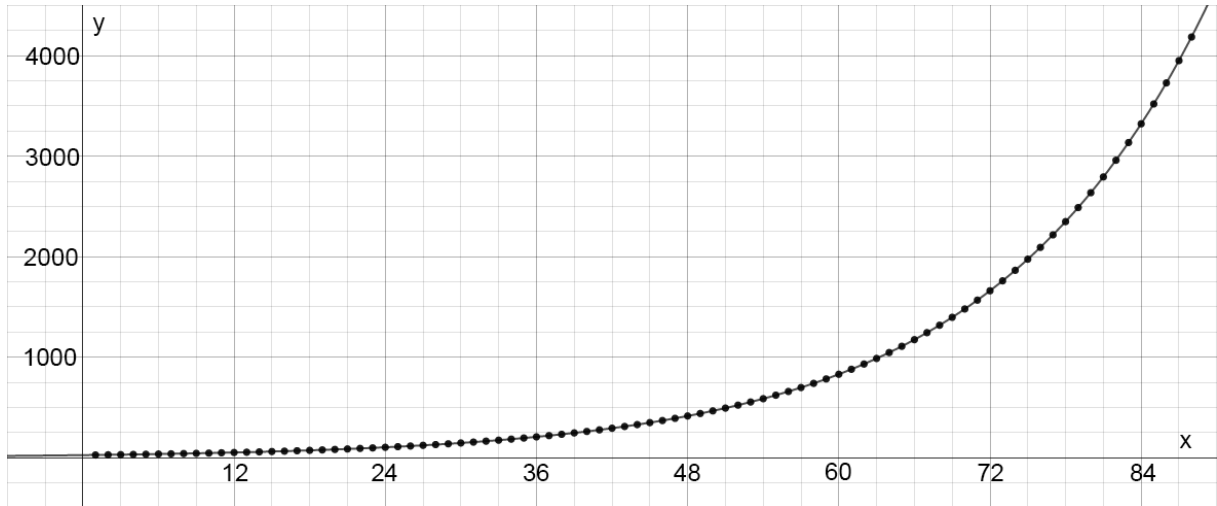


Figura 2.1: Frequência fundamental nas teclas do piano.

fixando-se a frequência f_k da tecla $N = k$. É comum que seja usada como referência a altura $f_{49} = 440\text{Hz}$. A Figura 2.1 apresenta o gráfico da função

$$y(x) = 440 \cdot 2^{\frac{x-49}{12}} \quad (2.2)$$

sobreposta pelos pontos descritos em (2.1) com $k = 49$ e $1 \leq N = x \leq 88$ para valores inteiros de x . O comportamento assintótico de (2.2) é um reflexo de como o crescimento exponencial da frequência fundamental é percebido pelo ouvido humano como um aumento linear da altura. Esta característica tem repercussões importantes no desenvolvimento de modelos físicos de instrumentos musicais, as quais serão apresentadas e discutidas no decorrer deste trabalho.

Capítulo 3

Guias Digitais de Ondas

3.1 Introdução

Conforme apresentado no capítulo anterior, a definição dos intervalos de uma escala musical é baseada num processo psicoacústico quanto à “agradabilidade” daqueles ao ouvido humano. No entanto, processos desta natureza fogem do escopo do presente trabalho. O que deve ser analisado são as características reticentes dos sons musicais considerados afinados, presentes e descritas na bibliografia da área. Para isso, mostra-se prudente que, dados os recursos disponíveis, sejam adotados modelos computacionais o mais próximos quanto possível de sistemas físicos já existentes, conhecidos e estudados para a produção destes sons, primando pela capacidade de replicação e mensuração. No contexto atual foi escolhido o modelamento de ondas transversais numa corda unidimensional, dada a sua simplicidade e disponibilidade de interpretação dos resultados devido à sua simplicidade.

Neste capítulo são interpretadas características do sistema real, que posteriormente serão levadas para o domínio teórico ideal. Quanto à implementação, são discutidas duas técnicas de modelamento físico: o algoritmo Karplus-Strong [5] e as Guias Digitais de Ondas [6].

3.2 A Corda Vibrante

Uma corda vibrante é um dos cenários mais simples de análise das propriedades de propagação de ondas, em especial devido às suas dimensões físicas: dado que a diferença nas ordens de grandeza entre suas medidas permite que sua seção transversal seja desprezível em relação ao seu comprimento, um sistema unidimensional é satisfatório como um ponto de partida [5, 6, 8, 10]. Colunas de ar em instrumentos cilíndricos também são considerados sistemas simples, e podem ser generalizados de maneira muito semelhante [6, 9].

Sabe-se que a propagação de ondas numa corda possui características transversais e longitudinais [20]. Enquanto as ondas transversais estão mais ligadas à altura e timbre, ondas longitudinais conferem detalhes de dispersão, atenuação e conceitos mais abstratos como “forma” e “leveza”. Estes conceitos serão retomados de maneira breve, visto que são importantes na produção musical em instrumentos reais, mas contribuem pouco para o objetivo deste trabalho. Foi adotada a estratégia de simular todas estas características apenas com a explicitação de ondas transversais contidas em um único plano, emulando

nelas, através da utilização de filtros digitais, os efeitos causados por ondas longitudinais. Como consequência, obtém-se um modelo bastante simplificado, e que atende o requisito proposto quanto à afinação.

3.2.1 Equação da Corda Ideal

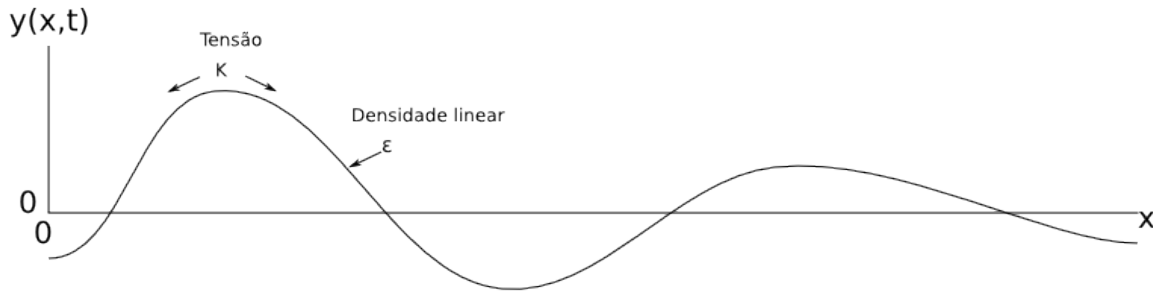


Figura 3.1: A corda vibrante ideal.

Consideremos uma corda elástica unidimensional, com densidade linear ε constante, que repousa ao longo da dimensão x sob uma tensão K , também constante. Caso um ponto arbitrário da corda seja deslocado, surgirão forças que tendem a restaurar a corda à sua posição inicial. De maneira concorrente, a inércia da corda irá retardar o retorno imediato à posição de equilíbrio, causando dois fenômenos importantes: a força restauradora que acelera a corda irá conferir momento às suas porções infinitesimais, as quais irão ultrapassar a posição de equilíbrio, “espelhando” o fenômeno; e a perturbação, que inicialmente era apenas local, deverá se propagar ao longo da corda com o passar do tempo. Tais fenômenos são explicados através das leis de conservação da energia e da transferência de momento [21].

De maneira a simplificar a análise, serão impostas algumas restrições a este modelo. Definimos que $y(x, t)$ descreve o deslocamento da corda em função de sua posição x e do tempo t . Todo deslocamento da corda a ser considerado estará contido no plano xy . Além disso, assumimos que o deslocamento em qualquer ponto da corda é pequeno o suficiente para que a magnitude da tensão seja considerada constante, o ângulo de inclinação da corda deslocada em relação ao eixo x é pequeno, e o deslocamento de uma porção infinitesimal da corda pode ser considerado como perpendicular (ou transversal) ao eixo x . Efeitos de dissipação de energia, como atrito, resistência do ar, rigidez da corda e força da gravidade serão desprezados.

Seja Δx um segmento infinitesimal da corda. Ao ser deslocado, este segmento sofrerá a ação da tensão em direções opostas, que por sua vez exercerá forças proporcionais ao seu ângulo nas extremidades em relação ao eixo x , aqui notados como θ_1 e θ_2 . É importante salientar que os ângulos são ligeiramente diferentes [21], o que resultará na força restauradora no sentido de retorno à posição de equilíbrio. $K \sin(\theta)$ e $K \cos(\theta)$ correspondem, respectivamente, às componentes vertical e horizontal da tensão nos pontos de fronteira, de maneira que a força resultante em Δx terá componentes $K(\sin(\theta_2) - \sin(\theta_1))$ na direção y e $K(\cos(\theta_2) - \cos(\theta_1))$ na direção x . Como θ_1 e θ_2 são pequenos, podemos

considerar que

$$\begin{aligned} \text{sen}(\theta) &\approx \text{tan}(\theta) = \frac{\partial y}{\partial x} \\ \text{cos}(\theta_1) &\approx \text{cos}(\theta_2), \end{aligned}$$

onde y representa o deslocamento médio de Δx . Desta forma, as componentes horizontais das forças exercidas serão desprezadas, e o comprimento de Δx será aproximado à sua projeção no eixo x .

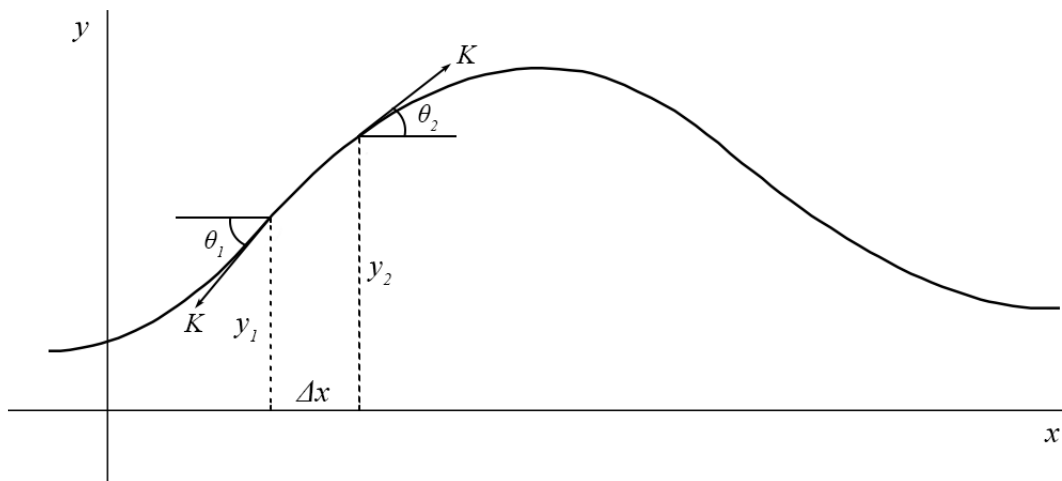


Figura 3.2: Forças atuantes num segmento da corda vibrante.

Aplicando a Segunda Lei de Newton, temos que a força resultante no segmento Δx deve ser igual à sua massa vezes sua aceleração. Logo, temos que

$$K \left(\frac{\partial y_2}{\partial x} - \frac{\partial y_1}{\partial x} \right) = \varepsilon \Delta x \frac{\partial^2 y}{\partial t^2}. \quad (3.1)$$

À medida em que Δx tende a zero, o deslocamento médio aproxima-se do deslocamento propriamente dito. Além disso, podemos dividir os dois lados de (3.1) por Δx , resultando em

$$K \frac{\partial^2 y}{\partial x^2} = \varepsilon \frac{\partial^2 y}{\partial t^2}. \quad (3.2)$$

Como resultado das restrições impostas, percebe-se que $y(x, t)$ é função da posição x e do tempo t , já que apenas o deslocamento transversal será considerado. Logo, faz-se pertinente as seguintes definições:

$$\dot{y} \equiv \frac{\partial}{\partial t} y(x, t) \quad (3.3)$$

$$y' \equiv \frac{\partial}{\partial x} y(x, t). \quad (3.4)$$

A força restauradora relevante é justamente a tensão na corda vezes a sua curvatura (Ky''). Tal força é balanceada a todo momento pela inércia da corda, que é igual à sua

densidade linear vezes a aceleração transversal ($\varepsilon\ddot{y}$). Sendo assim, aplicando as definições anteriores em (3.2), obtemos a seguinte equação diferencial da corda:

$$Ky'' = \varepsilon\ddot{y}. \quad (3.5)$$

Modos de Vibração

Ao se considerar um tamanho finito para a corda, e determinar extremidades fixas, interpretados como deslocamento transversal nulo, observa-se a reflexão das ondas resultantes do deslocamento inserido. Através de processos construtivos e destrutivos resultantes das interferências destas chamadas ondas propagantes, observa-se a caracterização de ondas estacionárias, conforme observado na Figura 3.3.

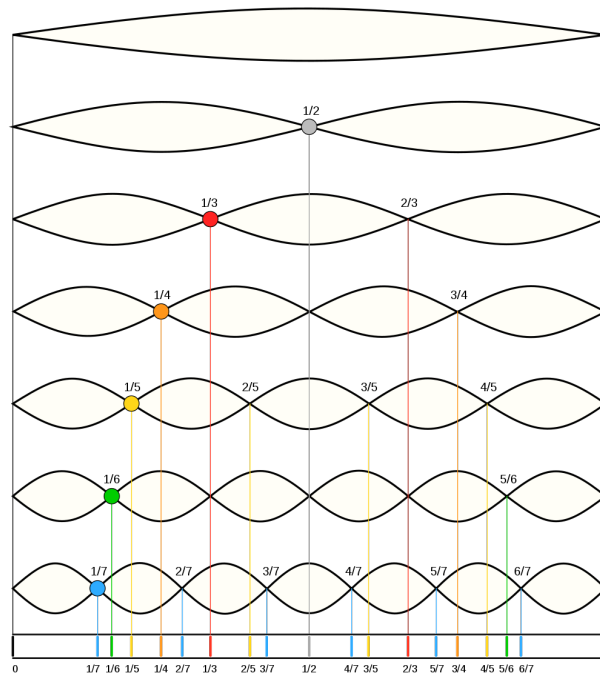


Figura 3.3: Modos de vibração de uma corda ideal.

Esta configuração corresponde aos diferentes modos de vibração da corda, que correspondem a frequências da série harmônica. Desta forma, uma corda ideal com extremidades fixas, quando deslocada, sempre irá vibrar em frequências harmônicas [21], resultando na produção de sons complexos com altura bem definida.

Equação de d'Alembert

D'Alembert propôs a substituição de variáveis na equação da onda unidimensional, de maneira que uma onda pudesse ser descrita por duas componentes que propagam em sentidos opostos, a saber

$$y(x, t) = y_r\left(t - \frac{x}{c}\right) + y_l\left(t + \frac{x}{c}\right). \quad (3.6)$$

A figura 3.4a apresenta um cenário em que a corda é tangida em três pontos simultâneos p , provocando um deslocamento triangular. Este deslocamento é interpretado como a

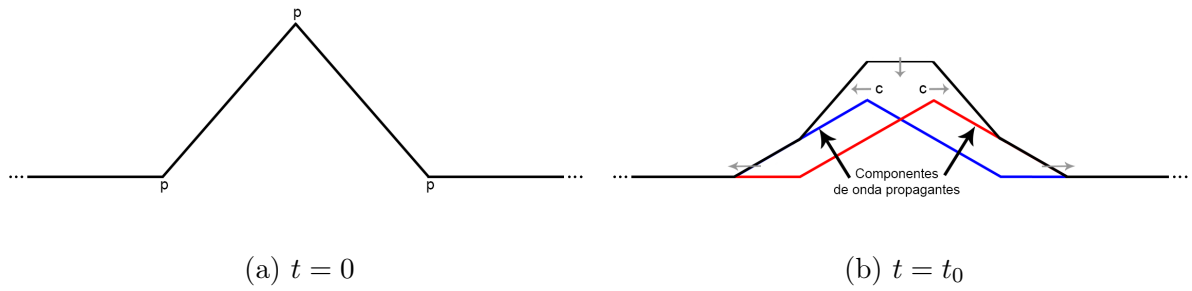


Figura 3.4: Visualização do teorema de d'Alembert.

sobreposição das duas componentes, cujas posições inicialmente coincidem. A figura 3.4b apresenta a mesma região da corda após um curto espaço de tempo t_0 , após o qual visualiza-se a decomposição do deslocamento à medida em que as componentes propagam em sentidos opostos com velocidade c .

Como estamos tratando de ondas unidimensionais numa corda elástica e sem perdas, num sistema causal e invariante no tempo, o deslocamento vertical de cada componente é transferido ao longo da corda de forma inalterada. Por ser um sistema linear, o deslocamento vertical da onda resultante é a simples soma de suas componentes. Ambas y_r e y_l são duplamente diferenciáveis, de maneira que elas respeitam as mesmas propriedades matemáticas da equação da corda (3.5).

Amostrando as Ondas Propagantes

Os passos para a amostragem das ondas já foram delineados anteriormente. A seguir, eles serão apresentados de forma explícita.

Tomemos novamente o intervalo temporal de amostragem $T = 1/f_s$, onde $f_s = 1/T$, e o intervalo espacial de amostragem $X = cT$. Devemos agora substituir as variáveis de domínio contínuo pelas de domínio discreto, da seguinte forma

$$\begin{aligned} x &\rightarrow x_m = mX, \\ t &\rightarrow t_n = nT. \end{aligned}$$

Aplicando esta substituição em (3.6), obtemos

$$y(x_m, t_n) = y_r((n - m)T) + y_l((n + m)T). \quad (3.7)$$

Já temos descritas em (3.7) as componentes de ondas já amostradas, cujas interferências resultarão na equação da corda vibrante unidimensional sem perdas. Percebe-se que o intervalo temporal de amostragem está presente em todos os parâmetros. Logo, para simplificação de notação, definimos

$$\begin{aligned} y^+[n] &\equiv y_r(nT) \\ y^-[n] &\equiv y_l(nT). \end{aligned}$$

Cada componente é simulada em uma linha de retardo. Neste arranjo, os pontos de excitação e captação podem ser arbitrariamente escolhidos para qualquer posição mX ao

longo da corda, oferecendo maior flexibilidade e, conseqüentemente, fidelidade de dinâmica e timbre.

3.3 Realizações

3.3.1 Guias Digitais de Ondas

Seguindo esta linha, Julius Smith cunhou a técnica de modelagem física de instrumentos musicais utilizando Guias Digitais de Ondas (*Digital Waveguides*) [6]. Trata-se de uma generalização do algoritmo Karplus-Strong [5], que pode ser estendido para sistemas n -dimensionais. Isto foi alcançado ao se explicitar a solução de ondas propagantes da equação de d'Alembert através de linhas de retardo distintas. Desta forma, ganha-se em flexibilidade quanto à observação de interferências entre as ondas. A seguir será apresentada a aplicação deste método no sistema unidimensional presente.

Topologia

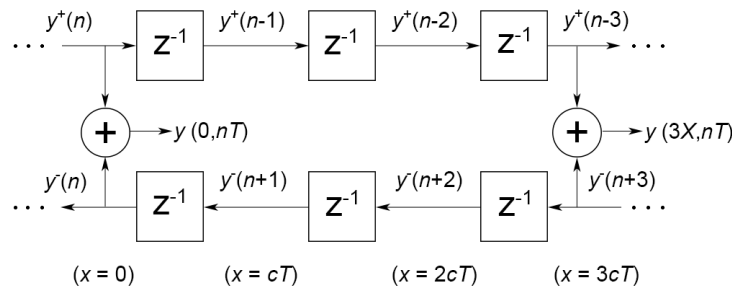


Figura 3.5: Topologia do algoritmo de Guias de Ondas Digitais.

Nas Guias de Ondas digitais, as linhas de retardo podem ser desmembradas até a menor parte factível, que é o retardo unitário, aqui denotado por z^{-1} . Conforme discutido anteriormente, o retardo unitário corresponde a um retardo de período temporal T ou, numa analogia a um instrumento real, a um intervalo espacial X percorrido por uma onda de velocidade c num período T . Esse desmembramento permite que características variáveis de dispersão e atenuação ao longo do corpo do instrumento sejam distribuídas no modelo. É comum que estas características sejam aproximadas através de multiplicadores lineares, ou filtros digitais análogos ao filtro de laço presente no algoritmo Karplus-Strong.

Correspondência com fenômenos físicos

A grande flexibilidade das Guias Digitais de Ondas oferece novas possibilidades, como a seleção de pontos de excitação e de observação. Além disso, devido à natureza linear e causal do sistema, fatores de perda podem ser consolidados em pontos arbitrários, assim como os próprios blocos de retardo, simplificando a implementação do modelo.

Uma das condições presentes numa corda com extremidades fixas é a de deslocamento zero em suas extremidades. Ao contrário do algoritmo Karplus-Strong, que apresenta esta

característica de forma implícita, na técnica de Guias Digitais de Ondas ela é apresentada explicitamente na modelagem do processo de reflexão.

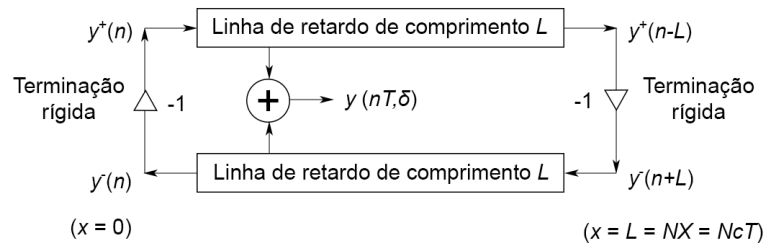


Figura 3.6: Guia de Ondas Digitais com linhas de retardo consolidadas e terminações reflexivas.

Sempre que uma das componentes de onda propagantes chega à terminação fixa, ela é refletida através da mudança de fase em πrad , obtida pela multiplicação do sinal por -1 . Posteriormente, este sinal é entregue à outra linha de retardo, simulando a sua propagação na direção oposta.

O resultado sonoro da simulação pode ser obtido a partir da soma das amostras presentes em pontos equivalentes das duas linhas de retardo, caracterizando em um ponto análogo à posição em um instrumento físico. De maneira semelhante, a excitação é introduzida no sistema seguindo a mesma analogia.

A principal característica dos sons produzidos pelo método de Guias Digitais de Ondas é a harmonicidade, ou seja, a produção de sons harmônicos, com altura bem definida. O comprimento L das linhas de retardo utilizadas definem a frequência fundamental f do som produzido como $f = f_s/L Hz$, onde f_s é a frequência de amostragem utilizada. A Figura 3.8 apresenta o espectro de um som produzido por esta técnica.

3.3.2 Algoritmo Karplus-Strong

O algoritmo Karplus-Strong foi uma das primeiras propostas que associaram uma linha de retardo com o comprimento de uma corda vibrante, constituindo o pioneiro dos modelos de síntese de sons por modelamento físico de instrumentos [5].

Um sinal amostrado $x[n]$, com larga banda de frequência, é oferecido à saída $y[n]$ ao mesmo tempo em que preenche uma linha de retardo de L amostras. A partir da amostra $x[n+L]$, observa-se a interferência com a amostra $y[n]$, correspondente à amostra $x[n]$ atrasada em L amostras e atenuada pelo filtro de laço. Este sinal, da mesma forma, é oferecido à saída e reintroduzido na linha de retardo.

Observa-se que o som resultante depende do sinal de entrada, do comprimento da linha de retardo e do filtro de laço. Em linhas gerais, o comprimento da linha de retardo está intimamente ligado com a altura do som produzido: dada uma frequência de amostragem f_s , uma linha de retardo de L amostras irá produzir um som com frequência fundamental de $f_s/L Hz$, de maneira semelhante ao que se observa na técnica de Guias Digitais de Ondas. A natureza do sinal de entrada e a definição do filtro de laço interferem em seu timbre.

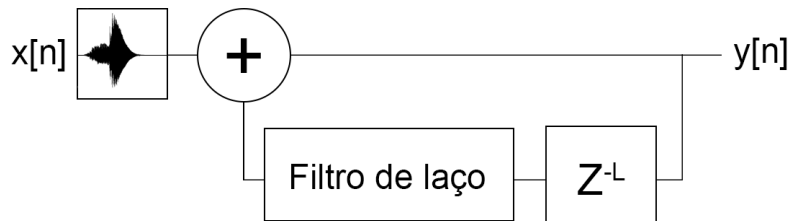


Figura 3.7: Esquema do algoritmo Karplus-Strong.

O algoritmo Karplus-Strong consiste numa generalização da técnica de Guias Digitais de Ondas, apresentando características bastante semelhantes. A Figura 3.9 apresenta o espectro de um som produzido seguindo o algoritmo Karplus-Strong, o qual pode ser comparado com o espectro presente na Figura 3.8. Note-se a predominância de espaçamento constante entre os vários valores de frequência, configurando a característica harmônica dos sons apresentados.

Podemos dividir o algoritmo em três eventos principais: **excitação**, **atenuação** e **cálculo de interferências**.

Excitação

A excitação consiste numa curta rajada de ruído branco, que é inserida no sistema. É comum que o ruído passe inicialmente por um filtro passa-baixas, a fim de evitar um acúmulo exagerado de energia no sinal de saída. Outras formas de excitação que podem substituir o ruído branco consistem num período de onda quadrada ou dente de serra, ou qualquer outra forma de onda com grande banda de frequência. A forma da excitação terá impacto direto no timbre do som produzido.

Atenuação

Numa corda real, a energia da excitação é perdida devido à sua natureza resistiva, devido ao atrito entre suas partículas e na interação com o ar. Estas perdas são concentradas no filtro de laço, que consolida e simula as perdas que uma onda propagante sofre entre as terminações da corda. É importante que este filtro provoque uma atenuação no sinal, a fim de, novamente, evitar o acúmulo exagerado de energia no sinal, e manter o sistema estável.

Cálculo de Interferências

Conforme o que se observa numa corda, a interferência causada por uma excitação pode ser descrita pela simples soma com o sinal presente no sistema. O algoritmo implementa esta característica com apenas um somador, demonstrando a sua simplicidade computacional.

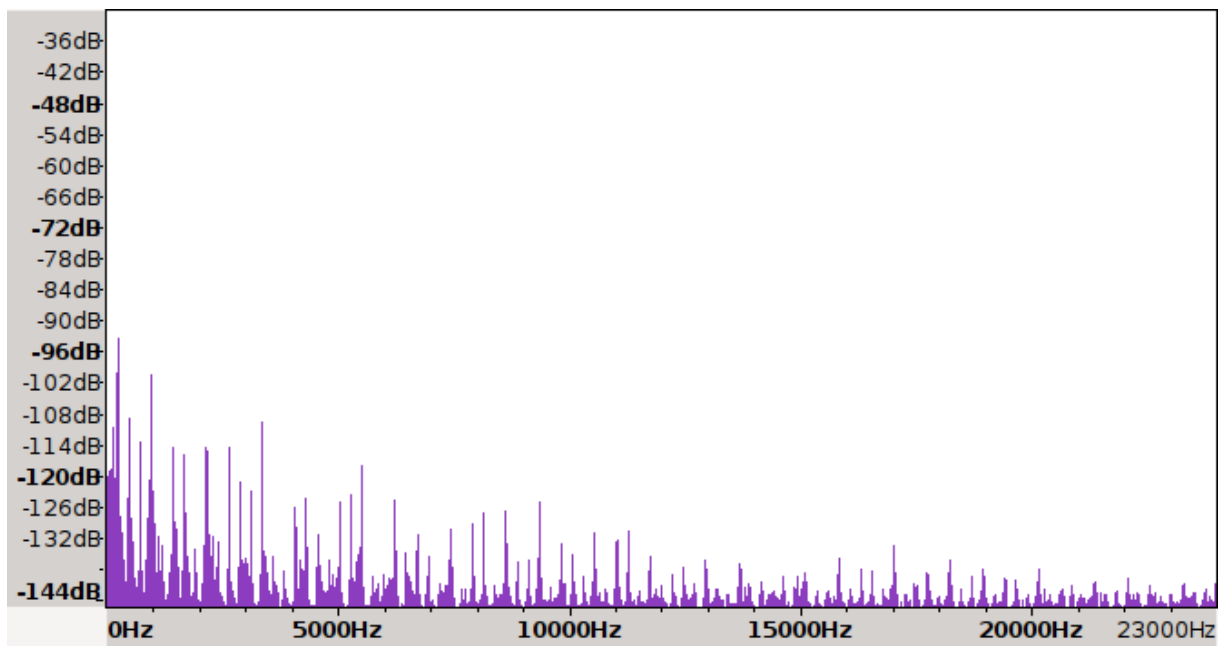


Figura 3.8: Espectro de um som com frequência fundamental de 440Hz , produzido pela técnica de Guias Digitais de Ondas.

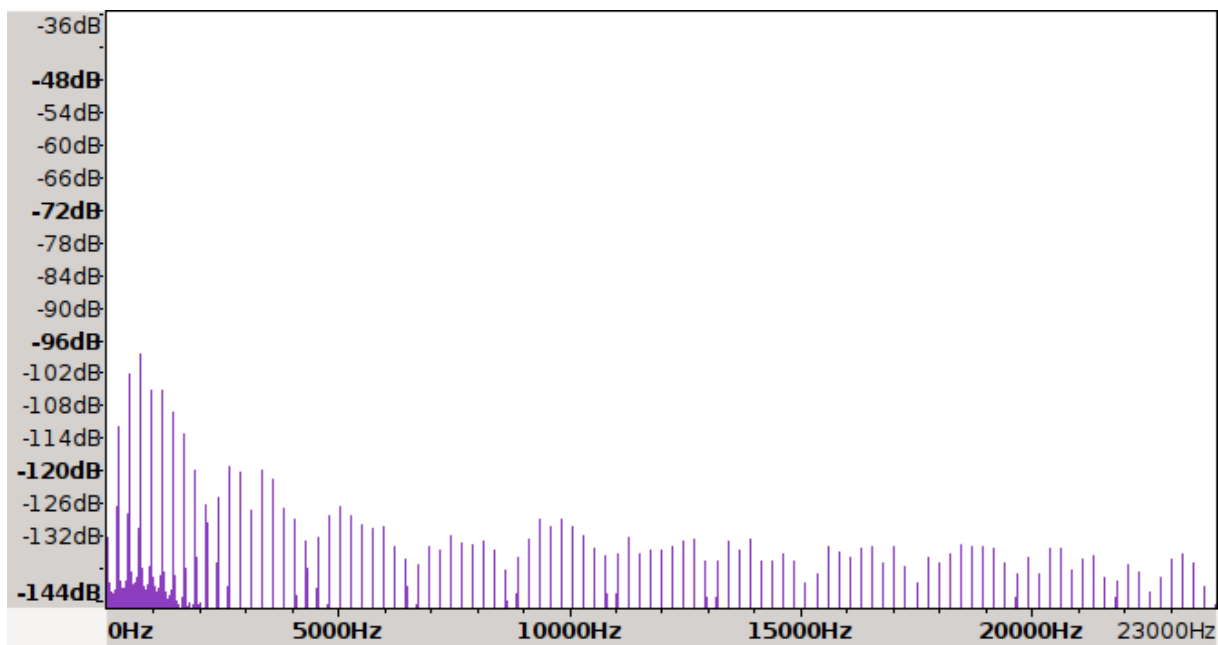


Figura 3.9: Espectro de um som com frequência fundamental de 440Hz , produzido pelo algoritmo Karplus-Strong.

Capítulo 4

Calibrando a Linha de Retardo

4.1 Introdução

Sabe-se que as guias digitais de ondas produzem uma frequência fundamental f_N respeitando a equação

$$f_N = f_s/N \quad [6], \quad (4.1)$$

onde f_s é a frequência de amostragem e N o comprimento da linha de retardo. No entanto, por se tratar de uma técnica de síntese digital, existe a limitação de que o tamanho das linhas de retardo deve ser inteiro. Tal característica está presente tanto no algoritmo Karplus-Strong quanto nas Guias Digitais de Ondas.

Esta limitação é crucial num sistema de síntese de sons musicais, pois pode impossibilitar a afinação das notas à determinada escala.

Tomemos como exemplo uma nota com frequência fundamental de $220Hz$, que corresponde à nota Lá mais próxima da nota Dó central do piano. Utilizando uma frequência de amostragem de $44100Hz$, que é a frequência de amostragem mais comum em sistemas digitais, e uma linha de retardo de comprimento $N = 200$ amostras, a frequência f_N produzida é de $220,5Hz$, enquanto que se aumentarmos a linha de retardo para 201, a frequência alcançada é, aproximadamente, $219,4Hz$. Ou seja, a menor alteração possível na linha de retardo, em uma amostra, nessa vizinhança gera uma mudança de mais de um Hz na frequência fundamental produzida.

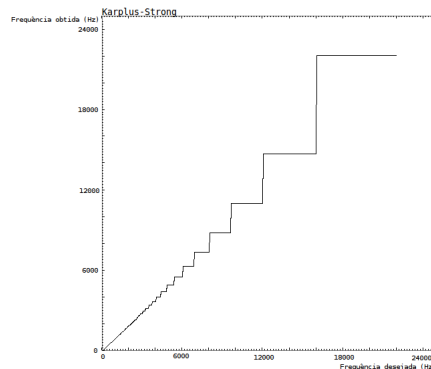


Figura 4.1: Frequências obtidas por uma linha de retardo de comprimento inteiro. Considerou-se uma frequência de amostragem de $48000Hz$.

Com o aumento da frequência desejada, e conseqüente diminuição da linha de retardo necessária, cada amostra torna-se crescentemente significativa. Em suma, a resolução de frequências diminui drasticamente à medida em que aproxima-se da frequência de Nyquist.

Ao retomarmos o exemplo anterior, percebe-se que para se obter uma frequência fundamental de 220Hz seria necessária uma linha de retardo de, aproximadamente, 200,45 amostras. A fim de se alcançar a afinação desejada, faz-se necessário empregar alguma técnica para aproximação de um retardo fracionário [16].

4.2 Filtros Digitais

O processo de retardo de um sinal pode ser interpretado como um processo de interpolação que segue os mesmos princípios da reamostragem de sinais, [16] sendo que a ferramenta mais usada para tal procedimento é o emprego de filtros digitais adequados. Várias são as técnicas utilizadas no design de filtros digitais para aproximação de retardos fracionários [9, 16, 20, 22–24]. A escolha do design a ser implementado depende do grau de fidelidade requerido, ou seja, da proximidade ao filtro de retardo ideal a ser alcançada.

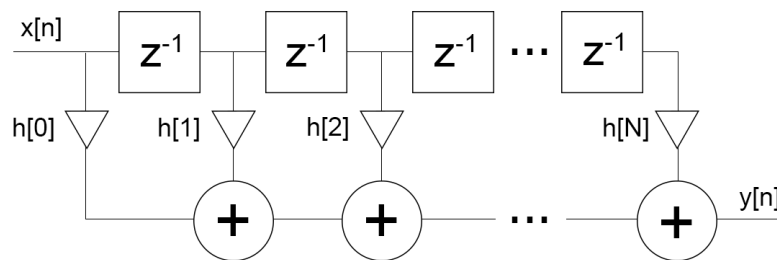


Figura 4.2: Realização de um filtro FIR de ordem N em sua Forma Direta.

O filtro de retardo ideal

A resposta em frequência de um filtro digital ideal que realiza um retardo D é dada por

$$H_{id}(e^{j\omega}) = e^{-j\omega D}, \quad |\omega| \leq \pi. \quad (4.2)$$

Percebe-se que sua magnitude é unitária, enquanto sua fase é dada por $-\omega D$, configurando um sistema *allpass* de fase linear [9].

O comportamento de um filtro digital¹ é observado através de sua resposta ao impulso unitário. Para valores inteiros de D , a resposta do filtro de retardo ideal consiste em uma única amostra não nula, correspondendo ao impulso atrasado em D amostras. No entanto, para atrasos de valores não inteiros, observa-se valores não-nulos para *todos* os pontos amostrais da função $\text{sinc}(\cdot)$, estendendo-se infinitamente antes e depois do ponto de referência temporal.

Tal característica impossibilita um comportamento causal ao sistema. De fato, foi constatado na literatura que isso impossibilita a própria realização de um filtro de retardo

¹As figuras Figura 4.3, Figura 4.4, Figura 4.5 e Figura 4.6 foram obtidas de [9].

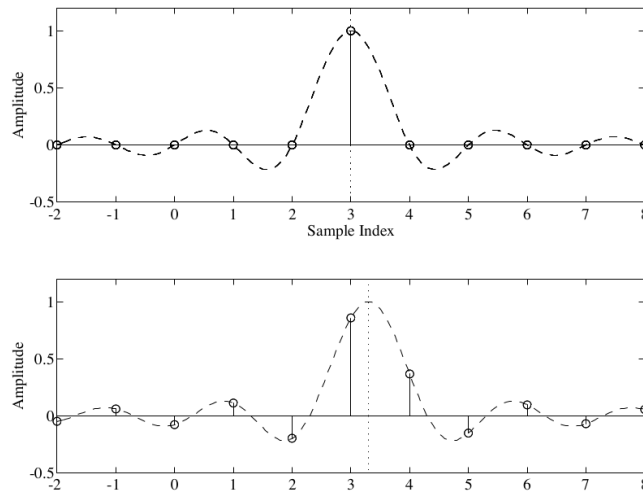


Figura 4.3: Função sinc atrasada em $D = 3$ e $D = 3.3$ amostras.

fracionário ideal [9]. Como este trabalho tem por objetivo apresentar uma implementação prática, é necessária a adoção de um modelo causal que *aproxime*, de maneira satisfatória, um valor fracionário de retardo.

4.2.1 Interpolador de Lagrange

O Interpolador de Lagrange consiste num filtro de resposta finita ao impulso (*FIR*) com a conveniente característica de oferecer planaridade máxima de frequência em torno da frequência zero. Tal característica é garantida ao se definir sua função de erro e as N primeiras derivadas como zero no nível DC. Desta forma, um filtro interpolador de Lagrange de ordem N apresenta uma resposta ao impulso com comprimento de N amostras [9, 16]. Planaridade máxima na frequência zero se encaixa perfeitamente com sistemas de sons musicais, visto que a audição humana percebe melhor sons graves que agudos.

Dado um valor de retardo fracionário D a ser aproximado pelo filtro, seus coeficientes podem ser determinados a partir da seguinte fórmula

$$h(n) = \prod_{\substack{k=0 \\ k \neq n}}^N \frac{D - k}{n - k}, \text{ para } n = 1, 2, \dots, N, \quad (4.3)$$

onde N representa a ordem do filtro.

Uma característica importante do interpolador de Lagrange é a *passividade*, que consiste numa resposta de magnitude menor ou igual a 1 em todas as frequências quando aproxima um delay fracionário [9]. Em sistemas que implementam retroalimentação, como é o caso dos algoritmos Karplus-Strong e Guias Digitais de Ondas, com a particularidade de não ser desejável a imposição de restrição de banda além da frequência de amostragem, esta característica é fundamental para a viabilidade da implementação. Qualquer ganho de magnitude maior que o unitário confere instabilidade ao sistema devido à possibilidade de acúmulo exagerado de energia no sinal.

É importante salientar que a passividade do filtro é restrita à aproximação de um retardo D contido num intervalo específico identificado pela sua ordem. [9, 23] Para filtros de ordem N ímpar, este intervalo encontra-se em

$$(N - 1)/2 \leq D \leq (N + 1)/2, \quad (4.4)$$

enquanto que para ordem par,

$$(N/2) - 1 \leq D \leq (N/2) + 1. \quad (4.5)$$

Aplicabilidade

Tomemos um retardo fracionário $D = \bar{D} + d$ a ser aproximado por um filtro interpolador de Lagrange, onde \bar{D} é inteiro e $0 \leq d \leq 1$ corresponde à parte fracionária. Percebe-se que filtros de ordem par são mais versáteis, pois são passivos numa amplitude de aproximação de 2 amostras, enquanto que nos filtros de ordem ímpar esta característica é restrita à amplitude de trabalho de 1 amostra. No entanto, a menor amplitude já é suficiente para que o filtro aproxime D qualquer que seja o valor de d em seu domínio.

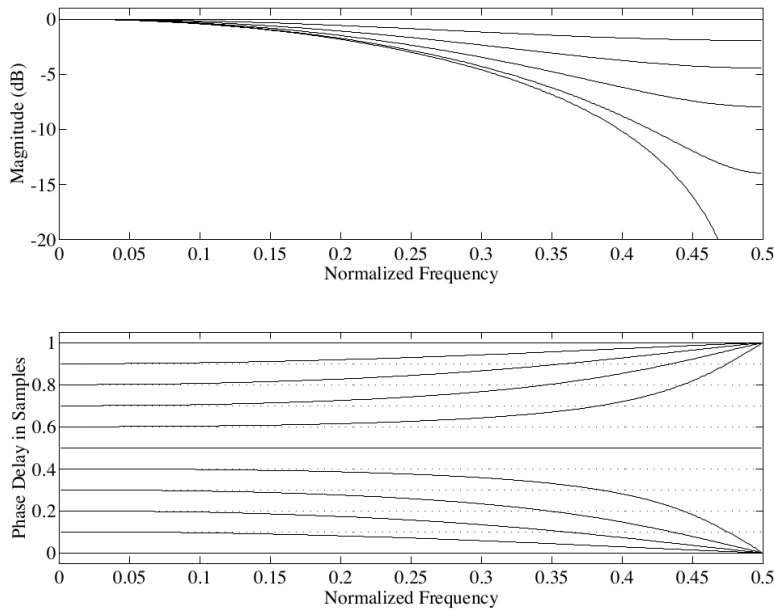


Figura 4.4: Interpolador de Lagrange de primeira ordem.

Além disso, para um mesmo N , o intervalo descrito em (4.4) está contido em (4.5). De fato, filtros de ordem N par apresentam erro mínimo de aproximação para valores contidos naquele intervalo [9], o que generaliza a definição do intervalo de trabalho dos filtros.

A paridade do filtro também apresenta características marcantes quanto à sua resposta de amplitude e linearidade de fase. Filtros de ordem ímpar serão sempre *notch filters* na vizinhança da frequência de Nyquist quando aproximam um valor $d = 0.5$ amostras, enquanto que filtros de ordem par apresentam atenuação mais suave para frequências mais altas. Em compensação, a aproximação do retardo de fase é mais próxima do ideal em filtros de ordem ímpar [16].

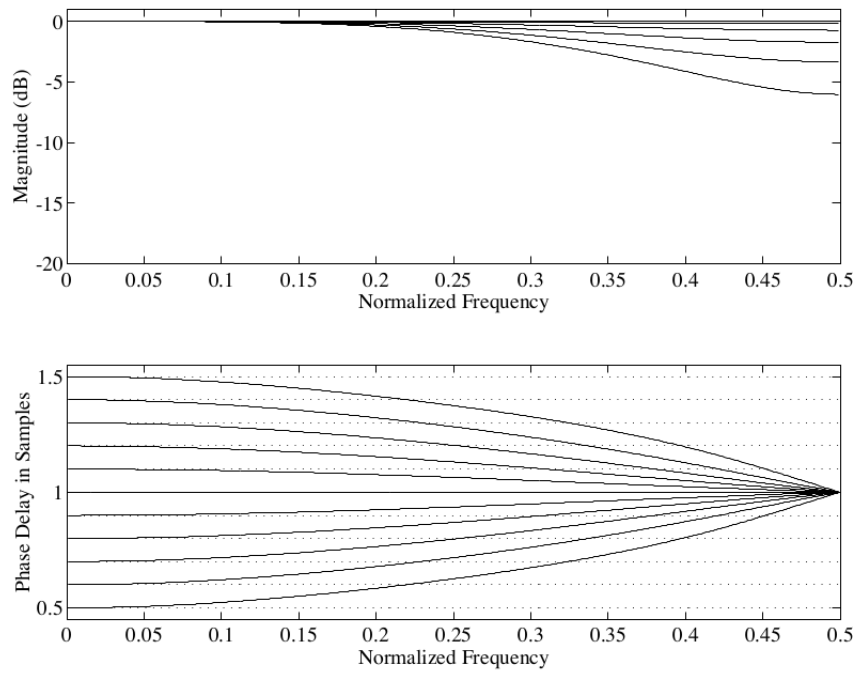


Figura 4.5: Interpolador de Lagrange de segunda ordem.

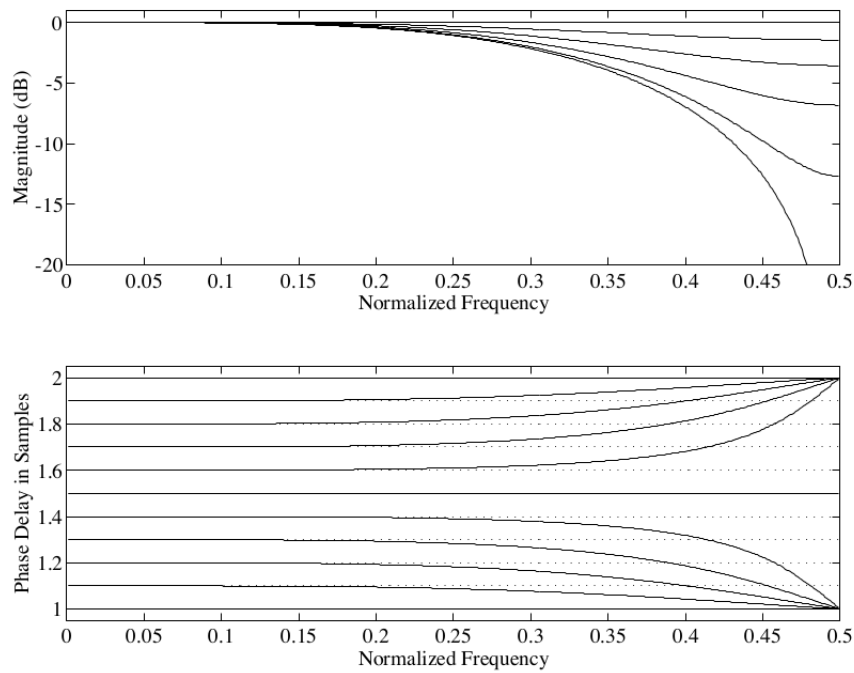


Figura 4.6: Interpolador de Lagrange de terceira ordem.

Capítulo 5

O Protótipo

5.1 Considerações iniciais

Conforme discutido no capítulo anterior, o uso de filtros de aproximação de retardo fracionário possibilitam uma resolução de frequência muito maior para sons produzidos utilizando métodos baseados ou derivados do algoritmo Karplus-Strong. De maneira geral, a aproximação melhora com o aumento da ordem do filtro, independente da paridade; e como qualquer aproximação assintótica não-linear, o aumento na complexidade computacional se traduz em ganhos comparativamente menores para altas ordens. Em contrapartida, o intervalo de trabalho do filtro também se desloca, acompanhando o aumento no número de blocos de retardo usados em sua realização, mas de forma linear.

Com efeito, limita-se a frequência fundamental máxima produzida pela linha de retardo, conforme observado em (5.3). De maneira a respeitar o limite de Nyquist, o valor mínimo de N deve ser 2, que corresponde a um retardo total (linha de retardo mais filtro de retardo fracionário) de duas amostras. Logo, para possibilitar a utilização de toda a banda de trabalho de uma determinada taxa de amostragem para produzir uma determinada frequência fundamental, a ordem máxima do filtro utilizado deve ser 3, conforme discutido em 4.2.1.

Estas constatações são importantes para a determinação do ambiente de trabalho que servirá como base para a implementação do protótipo, a começar pela determinação da frequência de amostragem.

Escolhendo a frequência de amostragem

Um dos instrumentos mais versáteis e difundidos no mundo é o piano de 88 teclas. Sua extensão engloba a de grande parte dos instrumentos existentes, produzindo alturas que vão de, aproximadamente, $30Hz$ a $4200Hz$, conforme apresentado na Figura 2.1. Por esta razão, é conveniente que o protótipo possa produzir, no mínimo, estas frequências fundamentais. No entanto, como diferentes componentes de frequência são responsáveis pelas características de timbre, e também porque podem mudar a percepção de altura de um som complexo, é desejável que sejam produzidos tantos harmônicos quanto possível.

Os primeiros processadores digitais de sinais foram criados num período em que tanto a memória quando o próprio poder computacional limitavam a escolha da frequência de amostragem. Felizmente, hoje a situação é diferente. É comum vermos frequências de

amostragem desde $8000Hz$ a $192000Hz$ para sinais sonoros. Como o ouvido humano dificilmente consegue discernir sons de frequência superior a $20kHz$, define-se um limite mínimo de $40kHz$ para a frequência de amostragem. O limite superior seria dado pelo poder computacional do sistema no qual o protótipo será executado.

Primando pela economia de recursos, e visando maior maleabilidade de replicação em diferentes sistemas, optou-se pela frequência de amostragem de $48000Hz$. Esta é a amostragem usada em sons digitais de sistemas modernos de DVD e Blu-ray, por exemplo. Uma frequência de amostragem mais alta traria benefícios quanto à resolução de frequências como um todo, mas também mascararia o desempenho dos filtros de retardo fracionário utilizados.

5.1.1 Modelo

O diagrama de blocos da Figura 5.1 a seguir representa o modelo do protótipo implementado. Trata-se de um modelo simplificado do algoritmo de Guias Digitais de Ondas, num caso que se reduz ao algoritmo Karplus-Strong: as linhas de retardo foram consolidadas, de maneira que os pontos de reflexão podem ser suprimidos. A aproximação das dispersões ocorridas durante a propagação das ondas é realizada através da multiplicação de cada amostra por um valor de amortecimento ψ no intervalo $0 \leq \psi \leq 1$. Esta operação funciona como um atenuador para todas as frequências presentes no sinal de forma equivalente. Além disso, o filtro de aproximação de retardos fracionários também atua como um atenuador para altas frequências, visto que ele tem, em geral, um comportamento passa baixas.

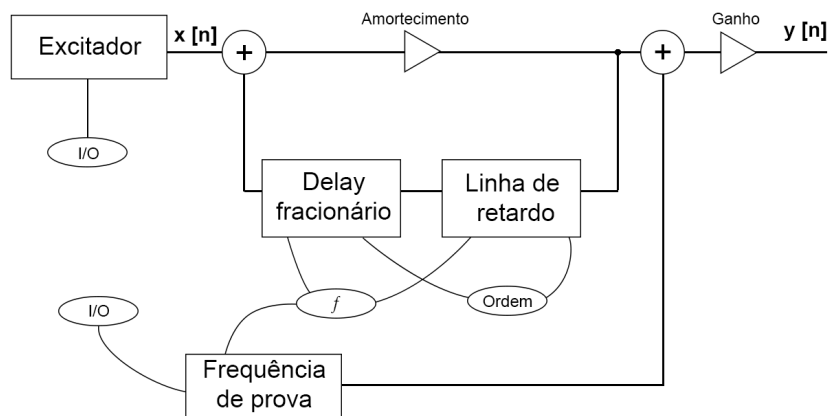


Figura 5.1: Diagrama-modelo implementado no protótipo.

O filtro pode ser escolhido entre interpoladores de Lagrange de primeira a terceira ordem, ou ignorado caso seja selecionada ordem 0. Neste caso, o arredondamento será feito para o inteiro mais próximo. Um oscilador senoidal pode ser habilitado para a produção de uma frequência de prova, para fins de comparação com as alturas produzidas pela linha de retardo. Este sinal é adicionado logo depois do laço da linha de retardo.

5.2 Implementação

O objetivo do presente trabalho consiste na obtenção de sons musicais afinados a uma escala musical arbitrária. O método utilizado consiste na realização de um modelo de síntese de áudio por modelamento físico de instrumentos musicais, num misto entre as técnicas de Guias Digitais de Ondas e do algoritmo Karplus-Strong. Não se trata, portanto, de uma nova técnica de síntese, mas sim da análise e, principalmente, experimentação de técnicas amplamente utilizadas e estudadas no meio acadêmico para a produção de sons musicais. A seguir serão discutidos os componentes implementados no protótipo do sintetizador.

Linguagem de programação

Dentre as várias opções disponíveis para modelamento e síntese de sons musicais, uma delas merece destaque pela sua facilidade de manipulação e disponibilidade de apresentação, aliadas a versatilidade: trata-se da linguagem de programação visual *PureData*¹ (*PD*), criada por Miller Puckette.

A versão utilizada foi *0.43.4-extended*, em sistema Linux com driver de audio ALSA. Nas configurações do *PD* foram configurados os seguintes parâmetros:

- Frequência de Amostragem: $48000Hz$
- *Delay*: $50msec$
- Tamanho de bloco: 64 amostras
- Output: 2 canais

Símbolos com um prefixo \$ representam variáveis definidas dentro do protótipo, e são significativas precisamente no escopo do *PD*.

5.2.1 O excitador

A excitação consiste no percorrimento de um vetor de 240 posições, feito em 5 milissegundos. Desta forma, cada posição corresponde, numa frequência de amostragem de $48000Hz$, a exatamente uma amostra.

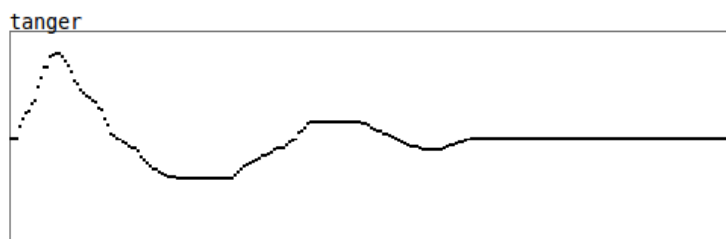


Figura 5.2: Esquema do Excitador. Cada amostra é oferecida diretamente à saída.

¹<http://www.puredata.info>

Este vetor pode ser facilmente editado, seja visualmente ou através de uma tabela. Esta característica confere ao sistema a facilidade de observação da influência que a excitação confere ao timbre do som produzido.

5.2.2 Linha de retardo

A linha de retardo foi realizada seguindo um misto entre o modelo Karplus-Strong e o de Guias Digitais de Ondas [5, 6]. Seu comprimento D é determinado pela equação (5.3) a partir de uma frequência alvo f previamente estabelecida, onde $N = D + d$, sendo D inteiro. O comprimento total necessário é recebido na variável $\$clr = N$. O valor de d será determinado de acordo com a ordem do filtro de retardo fracionário utilizado.

A entrada corresponde ao sinal proveniente do excitador, que é somado com a posição corrente da linha de retardo já atrasada pelo filtro de *delay* fracionário. Em seguida, sofre a ação do fator de amortecimento antes de ser realimentado na guia de ondas. Para ser entregue à saída, o sinal ainda é atenuado para evitar efeitos indesejados de *overflow*.

5.2.3 Afinação

A linha de retardo implementa uma parte inteira do retardo total necessário. A parte não-inteira, é definida de acordo com a ordem do filtro de retardo fracionário a ser utilizado. Filtros com ordens mais altas oferecem maior fidelidade na aproximação da parte fracionária do retardo necessário, mas limitam a frequência fundamental máxima de uma linha de retardo, conforme descrito em 3.3.2, 4.2.1 e discutido em 5.1. Foram implementados filtros interpoladores de Lagrange de primeira, segunda e terceira ordens, de modo a observar o efeito da ordem do filtro na fidelidade da afinação alcançada.

O filtro interpolador de Lagrange é realizado através da implementação de sua forma direta. O valor do retardo fracionário é recebido pela variável $\$fd$, e seus coeficientes são calculados conforme (4.3).

5.2.4 Usabilidade

A definição da frequência fundamental é feita através de uma frequência base e de deslocamentos de semitons ou comas. A frequência base pode ser definida através de um *slider* ou da entrada por teclado em uma caixa de edição, e fica disponível na variável $\$H0$. Os deslocamentos são obtidos através do clique nos botões *radio* correspondentes: os maiores correspondem aos $0 + 12$ semitons ζ presentes numa oitava, numa escala de temperamento igual, e os menores às $0 + 9$ divisões ς de um tom, chamadas de comas. Esta divisão foi alcançada através de testes empíricos, e representa o menor intervalo musical consistentemente perceptível pelo autor e por diversos estudantes com formação musical equiparável. A frequência fundamental $\$FF$ é obtida através da equação

$$\$FF = \$H0 \cdot 2^\kappa, \kappa = \frac{9\zeta + 2\varsigma}{108} \quad (5.1)$$

Depois de definida a frequência fundamental, um som de altura equivalente será produzido a cada clique no botão de entrada. Dois *sliders* são responsáveis por alterar linearmente o amortecimento e o ganho do sinal de saída.

5.3 Resultados

A análise dos resultados obtidos tem o objetivo de qualificar a *performance* dos diferentes filtros de aproximação de retardos fracionários quanto à afinação alcançada. Este trabalho foi idealizado para ouvintes humanos, de maneira que o requisito mínimo e satisfatório a ser alcançado é a do limiar da percepção de afinação pelo ouvido humano.

Além disso, conforme discutido no Capítulo 4, a própria natureza da síntese de áudio utilizando linhas de retardo garante maior resolução de frequências para frequências fundamentais graves. De fato, a utilização de filtros digitais para aproximação de retardos fracionários só tem efeito para linhas de retardo com comprimento menor que 290, visto que, para comprimentos maiores, a diferença sonora causada pela alteração da linha de retardo em uma amostra é audivelmente imperceptível.² Para a taxa de amostragem aqui utilizada, isso significa que, por construção, frequências menores que $48000/290 \approx 165,5Hz$ estarão sempre perceptivelmente afinadas.

A seguir são apresentados os resultados que corroboram com as informações apresentadas.

5.3.1 Metodologia

Todos os arquivos *wav* analisados constituem uma cópia direta da saída entregue ao conversor digital-analógico (*DAC*) da placa de som pelo *PD*. Trata-se de arquivos de canal único, com 240000 pontos amostrais – 5 segundos a uma taxa de amostragem de $48kHz$ – codificados como valores de ponto flutuante de 32 bits.

Cada arquivo foi analisado preliminarmente pelo programa de gravação e edição de áudio *Audacity*³ na versão 2.0.6, através da função *Plot Spectrum*. Esta função executou o algoritmo de FFT com os seguintes parâmetros:

- Algoritmo: *Spectrum*
- Função: janela de *Hanning*
- Tamanho da janela: 2^{16} amostras

Este tamanho de janela confere grande resolução de frequência, oferecendo intervalos de frequência (*bins*) de largura $\delta = f_s/2^{16} \approx 0,732422Hz$. Esta margem de erro pode mascarar diferenças audivelmente perceptíveis [17] na região grave, mas torna-se desprezível para frequências maiores que $210,966Hz$. Com efeito, o tamanho fixo de janela oferece precisão crescente com a frequência analisada. No entanto, sons produzidos com frequência fundamental na região entre $165,5Hz$ e $210,966Hz$ permanecem suscetíveis a mascaramento de imprecisões perceptíveis ao ouvido. Logo, a utilização de uma janela de menor comprimento mostrou-se necessária.

Como o maior tamanho de janela disponível no *Audacity* era justamente 2^{16} , optou-se pela utilização do programa *Octave* em sua versão 3.8.1. Nele é possível calcular a FFT com um tamanho arbitrário, aqui definido como 2^{17} . Desta forma, a precisão alcançada no

²Um ouvido bem treinado consegue perceber uma alteração mínima de 6 *cents*, correspondente a um intervalo musical de $1 : 2^{\frac{1}{200}}$ [17, 18]

³<http://audacity.sourceforge.net/>

espectro de frequência supera à do ouvido a partir de $105,483Hz$. A Figura X apresenta os gráficos das funções

$$\rho_1(f) = 1200 \cdot \log_2 \left(\frac{f + \delta'/2}{f - \delta'/2} \right), \delta' = f_s/2^{17} Hz \quad (5.2)$$

e

$$\rho_2(f) = 1200 \cdot \log_2 \left(1 + \frac{f}{f_s} \right) \quad (5.3)$$

de maneira sobreposta. O gráfico de (5.2), representado na cor azul, apresenta intervalos equivalentes à janela δ' que, quando observado nos pontos $N * \delta'$, $N = 1, 2, 3, \dots$, equivale à precisão de cada *bin* de frequência analisado pela FFT. O gráfico de (5.3), representado em vermelho, apresenta os intervalos em *cents* correspondentes à mudança em uma unidade no comprimento inteiro da linha de retardo em função da frequência a ser produzida, com destaque para frequências obtidas por tamanhos inteiros L da linha de retardo no intervalo de $2 \leq L \leq 290$. A Figura Y consiste numa ampliação da imagem anterior na região de interseção dos dois gráficos. Note-se a região com intervalos inferiores a 6 *cents*, indistinguíveis pelo ouvido.

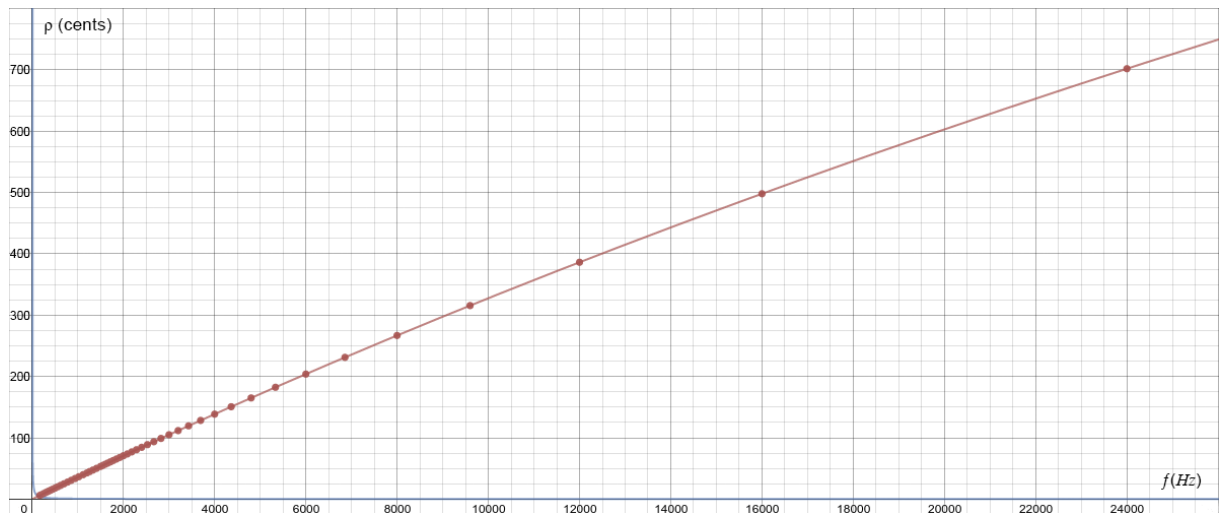


Figura 5.3: Precisão unitária da linha de retardo e binária da FFT.

A Tabela 5.1 apresenta a análise de sons produzidos pelo protótipo, com alturas equivalentes a algumas notas do piano. As 8 primeiras frequências alvo coincidem com a altura da primeira nota do piano e das próximas 7 distantes entre si por um intervalo de oitava. A penúltima frequência alvo representa a altura da última nota do piano. Foi feita, ainda, a extrapolação da extensão do instrumento para a altura correspondente a um intervalo de 8 oitavas da primeira nota, que está representada na última frequência alvo. *Bin* Máximo representa o *bin* de frequência com a maior magnitude na região da frequência alvo. $L+FD$ representa o tamanho inteiro L da linha de retardo utilizada, mais um valor fracionário de retardo aproximado pelo filtro. As linhas nas quais a ordem do filtro é 0 denotam situações em que não foi utilizado filtro para aproximação de retardo fracionário, para servir de base de comparação. Nestes casos, o tamanho da linha de

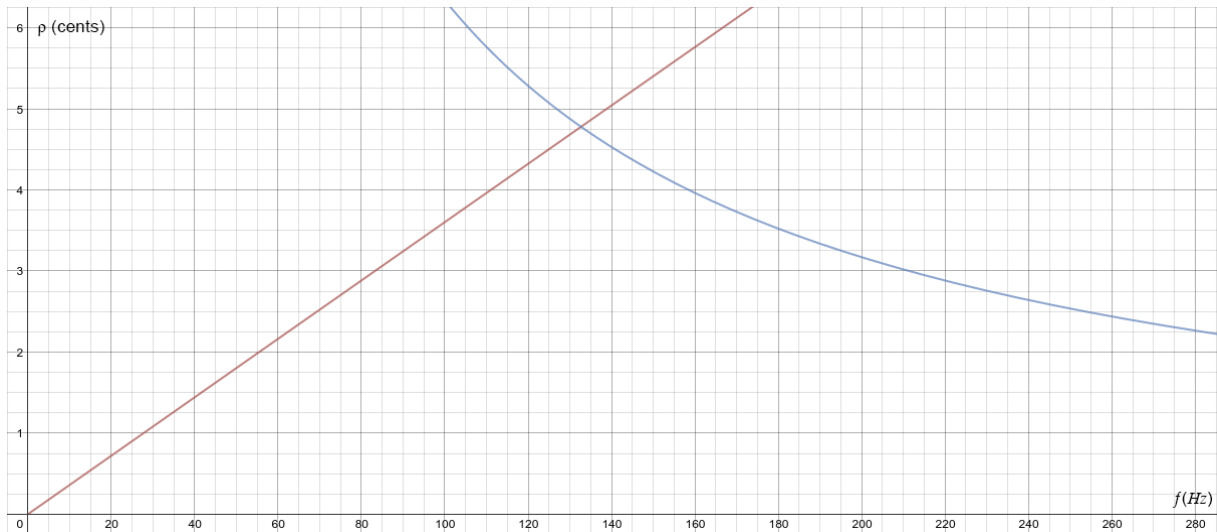


Figura 5.4: Detalhe da Figura 5.3.

retardo foi arredondado para o inteiro mais próximo. O desvio ρ foi calculado conforme

$$\rho = 1200 \cdot \log_2 \left(\frac{\text{Bin Máximo}}{\text{Frequência alvo}} \right), \quad (5.4)$$

com resultado em *cents*. Desvios de módulo inferior a 6 *cents* são imperceptíveis ao ouvido.

Conforme observado em Figura 4.4, Figura 4.5 e Figura 4.6, o erro máximo do interpolador de Lagrange corresponde a aproximações de retardos fracionários da forma $D + 0,5$, D inteiro. A Tabela 5.2 apresenta os desvios para estas aproximações em diferentes regiões de frequência.

5.3.2 Análise

Conforme discutido anteriormente, os filtros interpoladores de Lagrange têm eficiência máxima para aproximação de retardos da forma $d = 0$. Além disso, esperava-se que a eficiência apresentasse duas tendências: ela diminuiria à medida que d aproximasse de 0,5; e também diminuiria à medida que a frequência produzida aproximasse da frequência de Nyquist, ou seja, à medida que o comprimento da linha de retardo aproximasse de 0. Ambas as tendências foram observadas nos resultados apresentados, com exceção do desvio para frequência 7040Hz do filtro de primeira ordem, que ultrapassou o desvio do filtro de segunda ordem apenas neste ponto. Além disso, observa-se uma aparente anomalia ao se analisar o desvio do penúltimo valor de frequência presente na Tabela 5.1, referente à última nota do piano, a qual apresenta um valor discrepante na ausência de filtro de correção – ordem zero. Isto se deve ao fato de que o simples arredondamento da linha de retardo para o inteiro mais próximo não constitui a melhor aproximação possível, de modo a se obter o menor desvio da frequência alvo. No entanto, este problema não foi endereçado pela pouca relevância que oferece ao restante do trabalho: o menor desvio teórico seria alcançado com o arredondamento do valor fracionário que corresponda à

metade dos intervalos descritos em (5.3), ainda assim muito superiores ao observado pelos resultados com filtros de aproximação estudados.

Os resultados de análise espectral apresentados para aproximações de retardos da forma $N + 0.5$, N inteiro, presentes na Tabela 5.2 representam o máximo desvio possível nos intervalos de aproximação de cada filtro. Eles foram, em geral, satisfatórios, dados os requisitos iniciais do projeto. Mesmo o filtro de primeira ordem foi bastante eficaz na afinação percebida pelo ouvido até cerca de $500Hz$, a partir de onde a comparação simultânea com a frequência de prova denuncia a desafinação através do efeito de batimento. Apenas a partir de $1200Hz$ o desvio ficou perceptível ao ouvido, também devido ao efeito de batimento. Há que se observar que para frequências agudas uma desafinação de mais de 6 cents ultrapassa a distância máxima das frequências para que o efeito de batimento seja percebido pelo ouvido. Nesses casos, a análise de desvios de altura usando a FFT se faz necessária. O filtro de terceira ordem não apresentou desafinações perceptíveis em toda a extensão considerada, e apenas nas mais altas frequências o filtro de segunda ordem apresentou desvios perceptíveis. O caso anômalo do filtro de primeira ordem na produção de sons com frequência fundamental de $7040Hz$ não pôde ser analisado pelo ouvido, pois a forte atenuação causada pelo filtro impossibilita a sua compreensão de altura. O filtro de terceira ordem foi preciso mesmo na extrapolação da extensão do piano em quase uma oitava. Desta forma, o filtro interpolador de Lagrange de terceira ordem mostrou-se uma ótima ferramenta de afinação.

Apesar de se mostrarem, de forma quase unânime, sempre dentro da margem de erro considerada, houve casos em que foram percebidas ligeiras desafinações por alguns ouvintes expostos aos sons produzidos. Como o método de determinação da altura utilizado levou em conta apenas a frequência fundamental produzida, infere-se que os componentes de frequência superiores podem ter sofrido alterações em sua estrutura harmônica, por motivos não endereçados pelo autor, causando a desafinação percebida. A análise do impacto dos filtros apresentados na harmonicidade dos sons produzidos, bem como o emprego de técnicas mais elaboradas de determinação de altura de sons complexos, constituem empolgantes áreas para estudos futuros.

Frequência alvo	<i>Bin</i> Máximo	Ordem do filtro	$L+FD$	Desvio (<i>cents</i>)
27,50Hz	27,46582Hz	0	1745,00	$\approx -2,15$
	27,46582Hz	1	1745,45	$\approx -2,15$
	27,46582Hz	2	1745,45	$\approx -2,15$
	27,46582Hz	3	1745,45	$\approx -2,15$
55,00Hz	54,931641Hz	0	873,00	$\approx -2,15$
	54,931641Hz	1	872,73	$\approx -2,15$
	54,931641Hz	2	872,73	$\approx -2,15$
	54,931641Hz	3	872,73	$\approx -2,15$
110,00Hz	110,229492Hz	0	436,00	$\approx 3,61$
	109,863281Hz	1	436,36	$\approx -2,15$
	110,229492Hz	2	436,36	$\approx 3,61$
	110,229492Hz	3	436,36	$\approx 3,61$
220,00Hz	220,092773Hz	0	218,00	$\approx 0,73$
	220,092773Hz	1	218,18	$\approx 0,73$
	220,092773Hz	2	218,18	$\approx 0,73$
	220,092773Hz	3	218,18	$\approx 0,73$
440,00Hz	440,551758Hz	0	109,00	$\approx 2,17$
	440,185547Hz	1	109,09	$\approx 0,73$
	440,185547Hz	2	109,09	$\approx 0,73$
	440,185547Hz	3	109,09	$\approx 0,73$
880,00Hz	872,680664Hz	0	55,00	$\approx -14,46$
	880,004882Hz	1	54,55	$\approx 0,01$
	880,004882Hz	2	54,55	$\approx 0,01$
	880,004882Hz	3	54,55	$\approx 0,01$
1760,00Hz	1777,954101Hz	0	27,00	$\approx 17,57$
	1760,009766Hz	1	27,27	$\approx 0,01$
	1760,009766Hz	2	27,27	$\approx 0,01$
	1760,009766Hz	3	27,27	$\approx 0,01$
3520,00Hz	3428,466797Hz	0	14,00	$\approx -45,61$
	3518,920898Hz	1	13,64	$\approx -0,53$
	3517,089844Hz	2	13,64	$\approx -1,43$
	3520,019531Hz	3	13,64	$\approx 0,01$
4186,01Hz	4363,769531Hz	0	11,00	$\approx 72,00$
	4186,157226Hz	1	11,47	$\approx 0,06$
	4192,382813Hz	2	11,47	$\approx 2,63$
	4186,157226Hz	3	11,47	$\approx 0,06$
7040,00Hz	6857,299805Hz	0	7,00	$\approx -45,52$
	7002,685547Hz	1	6,82	$\approx -9,20$
	7015,50293Hz	2	6,82	$\approx -6,03$
	7037,109375Hz	3	6,82	$\approx -0,71$

Tabela 5.1: Picos de amplitude para sons de alturas presentes no piano e extrapolação da última altura.

$L+FD$	Ordem do filtro	Bin Máximo	Frequência alvo	Desvio (<i>cents</i>)
289,5	1	165,893555	165,80Hz	$\approx 0,94$
	2	165,893555	165,80Hz	$\approx 0,94$
	3	165,893555	165,80Hz	$\approx 0,94$
192,5	1	249,389648	249,35Hz	$\approx 0,27$
	2	249,389648	249,35Hz	$\approx 0,27$
	3	249,389648	249,35Hz	$\approx 0,27$
96,5	1	497,314453	497,41Hz	$\approx -0,33$
	2	497,314453	497,41Hz	$\approx -0,33$
	3	497,314453	497,41Hz	$\approx -0,33$
48,5	1	989,501953	989,69Hz	$\approx -0,33$
	2	989,868164	989,69Hz	$\approx 0,31$
	3	989,868164	989,69Hz	$\approx 0,31$
36,5	1	1315,063477	1315,07Hz	$\approx -0,007$
	2	1315,063477	1315,07Hz	$\approx -0,007$
	3	1315,063477	1315,07Hz	$\approx -0,007$
30,5	1	1573,608398	1573,77Hz	$\approx -0,18$
	2	1573,974609	1573,77Hz	$\approx 0,22$
	3	1573,608398	1573,77Hz	$\approx -0,18$
24,5	1	1959,228516	1959,18Hz	$\approx 0,04$
	2	1959,594727	1959,18Hz	$\approx 0,36$
	3	1959,228516	1959,18Hz	$\approx 0,04$
18,5	1	2594,604492	2594,59Hz	$\approx 0,007$
	2	2595,703125	2594,59Hz	$\approx 0,74$
	3	2594,604492	2594,59Hz	$\approx 0,007$
12,5	1	3839,72168	3840,00Hz	$\approx -0,12$
	2	3844,848633	3840,00Hz	$\approx 2,18$
	3	3840,087891	3840,00Hz	$\approx 0,04$
6,5	1	7382,812500	7384,62Hz	$\approx -1,80$
	2	7445,43457	7384,62Hz	$\approx 14,20$
	3	7384,277344	7384,62Hz	$\approx -0,08$

Tabela 5.2: Aproximações de retardos de 0.5 amostras, correspondentes ao erro máximo de aproximação.

Capítulo 6

Conclusão

Analisamos o processo de propagação de ondas numa corda unidimensional focando naquelas que produzem sons com altura determinável, mais especificamente os sons harmônicos naturalmente resultantes do cenário considerado. A partir disso, foi criado um modelo computacional para a síntese de sons musicais utilizando como base os modelos de Guias de Ondas Digitais e o algoritmo Karplus-Strong. A afinação dos sons produzidos foi alcançada através do emprego de filtros de resposta finita ao impulso que realizam a Interpolação de Lagrange. Foi discutido o impacto que diferentes ordens do filtro interpolador exercem no som produzido.

Através da metodologia e nível de fidelidade definidos no escopo do trabalho, concluiu-se que todos os filtros apresentam resultados satisfatórios para a extensão do piano, com destaque para o interpolador de Lagrange de terceira ordem. Ressalvas são feitas ao filtro de segunda ordem nas mais altas frequências analisadas.

No entanto, ainda que satisfatórios quanto à afinação, todos os filtros provocam alterações de timbre para aproximações de diferentes valores de retardos fracionários. Quando aproximam valores inteiros de retardo, todos se comportam como filtros de atraso ideais. No entanto, apresentam característica de passa-baixas para outros valores, sendo máximo também para aproximações de retardos da forma $N + 0.5$. Apesar de mais complexos computacionalmente, uma alternativa viável seria a utilização de filtros *all pass*, de modo a manter a uniformidade de timbre [16]. Destaca-se o filtro de Thiran, o qual apresenta características semelhantes às do interpolador de Lagrange, como a planaridade máxima no nível *DC*, por exemplo.

Deve-se destacar, também, o potencial do modelamento físico de instrumentos musicais como forma de desenvolvimento da produção musical. Seja para o aperfeiçoamento de instrumentos reais, seja para a criação de novos conceitos, a grande disponibilidade de síntese e teste de modelos matemáticos constitui uma ferramenta que alavanca a criatividade dos *luthiers*.

Referências

- [1] Curtis Roads. *The Computer Music Tutorial*. MIT Press, 1996. [1](#)
- [2] Charles Dodge and Thomas A. Jerse. *Computer Music - Synthesis, Composition and Performance*. Schirmer, 1997. [1](#)
- [3] G. Essl, S. Serafin, P. R. Cook, and J. O. Smith. Theory of banded waveguides. *Computer Music J.*, 28(1):37–50, 2004. [1](#)
- [4] L.R. Rabiner and B. Gold. *Theory and application of digital signal processing*. Prentice-Hall signal processing series. Prentice-Hall, 1975. [1](#)
- [5] K. Karplus and A. Strong. Digital synthesis of plucked-string and drum timbres. *Computer Music J.*, 7(2):43–55, 1983. [1](#), [2](#), [6](#), [11](#), [12](#), [23](#)
- [6] III Smith, Julius O. Physical modeling using digital waveguides. *Computer Music Journal*, 16(4):pp. 74–91, 1992. [1](#), [2](#), [6](#), [11](#), [15](#), [23](#)
- [7] Gianpaolo Borin, Giovanni De Poli, and Augusto Sarti. Algorithms and structures for synthesis using physical models. *Computer Music Journal*, 16(4):pp. 30–42, 1992. [1](#)
- [8] James Woodhouse. Physical modeling of bowed strings. *Computer Music Journal*, 16(4):pp. 43–56, 1992. [1](#), [6](#)
- [9] Vesa Välimäki. *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*. PhD thesis, 1995. [1](#), [2](#), [6](#), [16](#), [17](#), [18](#)
- [10] Giuseppe Cuzzucoli and Vincenzo Lombardo. A physical model of the classical guitar, including the player’s touch. *Comput. Music J.*, 23:52–69, June 1999. [1](#), [6](#)
- [11] Marc Aird and Joel Laird. Extending digital waveguides to include material modeling. In *In Proceedings of the Cost G-6 Conference on Digital Audio Effects (DAFX-01)*, pages 6–8, 2001. [1](#)
- [12] G K Yu, Y D Zhang, and Y Shen. Nonlinear amplitude-frequency response of a helmholtz resonator. *Journal of Vibration and Acoustics*, 133(2):024502–3, 2011. [1](#)
- [13] J. Makhoul. Linear prediction: A tutorial review. *Proceedings of the IEEE*, 63(4):561–580, April 1975. [1](#)
- [14] J.D. Markel and A.H. Gray. *Linear prediction of speech*. Communication and cybernetics. Springer-Verlag, 1976. [1](#)

- [15] J. Smith and P. Gossett. A flexible sampling-rate conversion method. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84.*, volume 9, pages 112–115, mar 1984. [1](#)
- [16] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine. Splitting the unit delay — Tools for fractional delay filter design. *IEEE Signal Processing Mag.*, 13(1):30–60, 1996. [2](#), [16](#), [17](#), [18](#), [30](#)
- [17] W. Apel. *Harvard Dictionary of Music*. Series I: Diaries. Belknap Press of Harvard University Press,c, 1969. [3](#), [24](#)
- [18] M.R. Jones, R.R. Fay, and A.N. Popper. *Music Perception*. Springer Handbook of Auditory Research, 36. Springer New York, 2010. [3](#), [4](#), [24](#)
- [19] A. Klapuri and M. Davy. *Signal Processing Methods for Music Transcription*. Springer Science+Business Media LLC, 2006. [3](#), [4](#)
- [20] David R. Rowland and Colin Pask. The missing wave momentum mystery. *American Journal of Physics*, 67(5):378–388, 1999. [6](#), [16](#)
- [21] W.C. Elmore, W.C. Elmore, and M.A. Heald. *Physics of Waves*. Dover Books on Physics Series. Dover Publications, 1969. [7](#), [9](#)
- [22] J. P. Thiran. Recursive digital filters with maximally flat group delay. *IEEE Trans. Circuits Syst.*, 18(6):659–664, 1971. [16](#)
- [23] Julius O. Smith. *Introduction to Digital Filter Theory*. William Kaufmann, Los Altos, CA, 04/1985 1985. [16](#), [18](#)
- [24] N.J. Loy. *An engineer's guide to FIR digital filters*. Prentice-Hall, 1988. [16](#)